



HAL
open science

Les effets de l'internat d'excellence de Sourdun sur les élèves bénéficiaires : résultats d'une expérience contrôlée

Luc Behaghel, Clément de Chaisemartin, Axelle Charpentier, Marc Gurgand

► To cite this version:

Luc Behaghel, Clément de Chaisemartin, Axelle Charpentier, Marc Gurgand. Les effets de l'internat d'excellence de Sourdun sur les élèves bénéficiaires : résultats d'une expérience contrôlée. [0] 2013. hal-02809550

HAL Id: hal-02809550

<https://hal.inrae.fr/hal-02809550>

Submitted on 6 Jun 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Les effets de l'internat d'excellence de Sourdun sur les élèves bénéficiaires : résultats d'une expérience contrôlée

Luc Behaghel*
Axelle Charpentier‡

Clément de Chaisemartin†
Marc Gurgand§

Rapport pour le Fonds d'expérimentation pour la jeunesse

Avec le concours de la DEPP¶

11 avril 2013

*École d'économie de Paris (INRA), J-Pal et IPP

†École d'économie de Paris et CREST

‡J-Pal

§École d'économie de Paris (CNRS), J-Pal et IPP

¶Direction de l'évaluation, de la prospective et de la performance du ministère de l'Éducation nationale. Nous sommes particulièrement reconnaissants envers M. Jean-François Chesné et Mme Sylvie Fumel qui ont conçu la batterie de tests de mathématiques et de français que nous avons utilisés pour mesurer les compétences des élèves, et envers M. Pascal Bessonneau et M. Thierry Rocher qui ont construit les indicateurs de compétence des élèves dans ces deux matières.

Remerciements

Cette recherche a été financée par le Fonds d'Expérimentation pour la Jeunesse, dans le cadre d'une convention avec l'Ecole d'économie de Paris et le rectorat de Créteil.

Nous remercions très vivement l'ensemble des partenaires qui nous ont aidés au cours des quatre années de ce projet. En premier lieu, nous tenons à remercier tout le personnel de Sourdun et plus particulièrement MM. Jean-François Bourdon, Bernard Lociciro, Stéphane Lours et Mme Sithi Maricar, qui ont tout mis en oeuvre pour que cette évaluation se déroule dans les meilleures conditions et la plus grande transparence. Nous voudrions aussi remercier la Direction de l'évaluation, de la prospective et de la performance (DEPP) du ministère de l'Éducation nationale : M. Jean-François Chesné et Mme Sylvie Fumel qui ont conçu la batterie de tests de mathématiques et de français que nous avons utilisés pour mesurer les compétences des élèves ; MM. Pascal Bessonneau et Thierry Rocher qui ont construit les indicateurs de compétence des élèves dans ces deux matières ; M. Cédric Afsa, Mme Pierrette Briant, M. Luc Brière, Mme Caroline Caroupapouille et Mme Chantal Payras qui nous ont fourni les données permettant de comparer les ressources déployées à Sourdun et dans les établissements des élèves témoins. Nous remercions aussi les membres de la Direction générale de l'enseignement scolaire (DGESCO) qui ont contribué au bon déroulement de cette évaluation, et plus particulièrement M. Jean Denis, Mme Catherine Marmiesse et Mme Bénédicte Robert.

Nous voulons ensuite remercier toutes les personnes du rectorat de Créteil impliquées d'une façon ou d'une autre dans cette étude pour leur aide précieuse : MM. Jean-Michel Blanquer, Jean Capillon, Jean-Thibault Daniel, Anthony Ernewein, Pascal Henri, William Marois, Cédric Montesinos, Thomas Renaud et Vincent Stanek. Nous remercions également la Mission d'animation du Fonds d'Expérimentation pour la Jeunesse, et notamment Mme Malika Kacimi et M. Mathieu Valdenaire.

Nous remercions très chaleureusement Adrien Bouguen qui a été chef de projet durant la première année et a mis en place l'ensemble du protocole d'enquête, ainsi que les nombreux assistants de recherche, en particulier Yann Guy, et le personnel administratif du laboratoire J-PAL au sein duquel la recherche a été menée, et qui ont effectué un travail considérable et de très grande qualité.

Enfin, nous exprimons notre profonde reconnaissance aux chefs d'établissement, aux élèves et à tout le personnel des 169 collèges et lycées que nous avons dû mettre à contribution pendant trois ans pour faire passer les questionnaires utilisés dans l'évaluation.

Les auteurs déclarent n'avoir aucun conflit d'intérêt en rapport avec les questions traitées dans cette recherche.

Synthèse

Les internats d'excellence constituent une orientation nouvelle et importante au sein des politiques scolaires. Ils visent à promouvoir la réussite d'élèves motivés, mais qui ne bénéficient pas d'un environnement social ou familial favorable pour développer leur potentiel. Le régime de l'internat doit permettre à ces jeunes de travailler dans de bonnes conditions et de bénéficier d'un encadrement adapté tout au long de la journée. On peut y voir une forme très intensive et volontariste de l'éducation prioritaire.

L'internat de Sourdun est le premier à avoir ouvert ses portes, à la rentrée 2009. Il a été suivi par 44 autres, si bien que 4 173 places étaient disponibles en internat d'excellence à la rentrée 2012. L'engagement budgétaire dont cette politique bénéficie, concentré sur un nombre relativement faible d'élèves, justifie les efforts mis en œuvre pour l'évaluer.

Une évaluation qui repose sur le tirage au sort des élèves admis à l'internat

Nous présentons ici les résultats d'une expérience contrôlée mesurant les effets de la scolarité à l'internat d'excellence de Sourdun. Cette expérience repose sur le tirage au sort des 258 élèves admis à l'internat en 2009 et 2010, parmi 395 candidats dont le dossier a été jugé conforme au profil attendu par une commission de sélection. Les 137 autres élèves n'ont pas pu être admis. Ils constituent alors notre groupe "témoin". Nous avons enquêté ces deux groupes d'élèves, ainsi que leurs familles, pendant les deux années suivant leur candidature. La situation des élèves témoins un an ou deux ans après le tirage au sort constitue le point de comparaison qui permet de mesurer l'effet de Sourdun sur les internes. En effet, le tirage au sort nous garantit que ces deux groupes étaient initialement identiques, et que les trajectoires des élèves témoins sont représentatives de celles qu'auraient connues les élèves internes s'ils n'avaient pas été admis à Sourdun.

Des élèves d'origine modeste et de niveau scolaire intermédiaire

Les élèves sont accueillis à l'internat de Sourdun à l'issue d'un double processus de sélection : candidature des familles – à l'initiative de celles-ci ou de l'établissement d'origine de l'élève – puis validation par une commission de recrutement. Un peu moins de la moitié d'entre eux sont boursiers, soit le double de la moyenne nationale. Un tiers d'entre eux appartiennent à une famille monoparentale, et plus de la moitié parle une autre langue que le français à la maison. Leurs parents sont plus nombreux à être au chômage que ceux des autres élèves de l'académie de Créteil.

En termes de niveau scolaire, ce sont d'assez bons élèves par rapport aux camarades de classe des établissements dont ils sont issus, mais des élèves intermédiaires comparés à la

moyenne nationale. Dans leurs classes d'origine, ils se situent typiquement au 7ème ou 8ème rang sur 25 élèves, en français et en mathématiques. Ceux qui ont passé le diplôme national du brevet (DNB) avant d'entrer à Sourduin n'ont presque jamais échoué (alors que le taux d'échec est de 20% au niveau national) et leurs notes les situent vers le 45ème rang sur 100 au niveau national. Il est donc important de garder à l'esprit que le dispositif dont nous évaluons l'impact s'adresse à ce type d'élèves qui, loin d'être les élèves excellents que certains internats attendaient peut-être, ne sont pas non plus des élèves faibles.

Au bout de deux ans, Sourduin produit un effet important sur les compétences des élèves en mathématiques

Nous avons tout d'abord quantifié l'impact de l'internat sur les compétences des élèves en français et en mathématiques, mesurées via des tests standardisés conçus par la Direction de l'évaluation, de la prospective et de la performance (DEPP) du ministère de l'Éducation nationale, et soumis dans les mêmes conditions à l'ensemble de l'échantillon d'élèves. Si, au bout d'un an, nous n'observons pas de différence entre les deux groupes, ni en français ni en mathématiques, au bout de deux ans les élèves internes obtiennent des résultats significativement plus élevés au test de mathématiques que les élèves témoins. La différence est de l'ordre de 41% d'un écart-type du score, ce qui signifie qu'en moyenne, un élève qui était initialement classé 45ème sur 100 atteint grâce à l'internat le même niveau que celui qui était classé 30ème. Dans les études sur les effets des politiques scolaires, un impact de cet ordre est considéré comme élevé; peu d'interventions permettent d'obtenir de tels progrès.

Ces résultats montrent tout d'abord que Sourduin produit des effets sur les compétences des élèves au bout de deux ans seulement. Il semble que ceux-ci passent d'abord par une phase de découragement lorsqu'ils arrivent à l'internat : la confiance des élèves en leurs propres capacités scolaires diminue lors de leur première année à Sourduin, avant de remonter. Ce processus est aussi décrit dans des travaux monographiques menés par ailleurs sur d'autres internats d'excellence.

D'autre part, l'absence d'impact en français peut surprendre. Il convient d'abord de souligner que ce résultat peut être lié à la taille limitée de notre échantillon, qui ne permet pas de détecter de différences entre les deux groupes lorsqu'elles sont petites. S'il existe des effets en français, mais qu'ils sont faibles, ils peuvent ne pas être apparents avec l'échantillon d'un peu moins de 400 élèves dont nous disposons. Soulignons aussi qu'il n'est pas rare, lorsque des évaluations rigoureuses d'interventions scolaires sont menées, d'obtenir des impacts forts en mathématiques et sensiblement plus faibles ou non significatifs en langue. C'est par exemple une tendance que l'on retrouve dans les évaluations des *charter schools* américaines, le dispositif le plus proche des internats d'excellence évalué dans la littérature internationale. Il

semblerait donc qu'il soit plus difficile d'accélérer les apprentissages en langue qu'en mathématiques, surtout chez des adolescents. Certains travaux en psychologie du développement montrent en effet que les aptitudes verbales des enfants sont fixées plus tôt dans la vie que leurs aptitudes numériques et mathématiques.

D'après nos calculs, la dépense par élève à Sourdun est environ deux fois supérieure à celle que l'on observe dans les établissements où sont scolarisés les élèves témoins, une différence qui s'explique en grande partie par le coût de l'internat. On peut donc comparer l'impact de Sourdun à celui que l'on obtiendrait si on divisait par deux la taille des classes, puisque cela conduirait également à doubler la dépense par élève. Les études disponibles montrent que cette politique aurait un impact sur les résultats scolaires des élèves similaire à celui de Sourdun. Si l'on s'en tient aux effets qu'il produit sur les compétences cognitives des élèves, l'internat d'excellence a donc un rapport coût-efficacité comparable à celui d'une réduction de moitié de la taille des classes. Faute de disposer d'études mesurant de façon rigoureuse les impacts d'autres interventions (comme par exemple les interventions ciblant les jeunes enfants), nous ne pouvons malheureusement pas mettre leurs effets en perspective avec ceux produits par Sourdun.

Sourdun a aussi un fort impact sur l'ambition des élèves et celle de leurs parents

Mais l'internat d'excellence a également des effets sur certaines dimensions non cognitives, dont de nombreuses études montrent qu'elles peuvent avoir des conséquences importantes sur la carrière scolaire et l'insertion sociale.

Nous avons ainsi mesuré l'ambition scolaire des élèves et celle que leurs parents ont pour eux. Il ressort que deux ans après le tirage au sort, les internes sont presque trois fois plus nombreux que les élèves témoins à déclarer souhaiter intégrer une classe préparatoire. Ils sont aussi 25% plus nombreux à dire qu'ils souhaitent obtenir un master. On retrouve cet effet chez leurs parents. En revanche, les parents d'un élève interne à Sourdun n'ont pas une ambition accrue pour ses frères et sœurs non-internes.

Nous avons aussi proposé aux élèves des tests permettant de mesurer leur motivation pour l'école. Parmi les sources de motivation identifiées par les psychologues, nous trouvons que les élèves scolarisés à l'internat de Sourdun ont, après deux ans, une plus grande motivation "intrinsèque" que les élèves témoins. Cela signifie qu'ils s'intéressent davantage aux études pour elles-mêmes, par opposition à une motivation qui verrait dans l'école un instrument pour atteindre d'autres buts.

Sourdun fournit aux élèves un cadre propice au travail

Afin de déterminer quels sont les facteurs susceptibles d'expliquer les impacts produits par l'internat, nous avons comparé la scolarité des élèves à Sourdun à celle des élèves témoins.

Notons tout d'abord que la charge de cours à Sourdun suit les recommandations des programmes officiels : les élèves n'y bénéficient donc pas de plus d'heures de français ou de mathématiques. En revanche, leur présence en classe est plus régulière : ils sont par exemple trois fois moins nombreux que les élèves témoins à déclarer sécher les cours. Les classes sont plus petites à Sourdun (19 élèves contre 24 en moyenne dans le groupe témoin). Les élèves y sont trois fois moins nombreux à dire que leurs camarades n'écoutent pas ce que disent les professeurs, et ils sont aussi deux fois moins nombreux à déclarer que des actes de violence sont fréquemment commis par d'autres élèves dans l'enceinte de l'établissement. Les heures de cours à Sourdun semblent donc se dérouler dans de meilleures conditions que dans les établissements témoins. Cela peut s'expliquer au moins en partie par le ciblage des élèves : l'internat s'affranchit de la carte scolaire, et regroupe ensemble des jeunes d'assez bon niveau et décidés à progresser.

Les professeurs de Sourdun ont été sélectionnés sur profil, notamment sur leur adhésion au projet éducatif. Cette procédure a conduit au recrutement de professeurs plus jeunes et plus qualifiés que ceux qui enseignent dans les établissements témoins : ils ont en moyenne trois ans de moins d'ancienneté et 28% d'entre eux sont agrégés contre 18% dans les établissements témoins. Le fait d'enseigner dans des classes relativement petites et homogènes leur permet de déployer plus d'efforts de pédagogie : par exemple, les élèves de Sourdun sont deux fois plus nombreux que les élèves témoins à déclarer que leurs professeurs continuent à expliquer jusqu'à ce que tous les élèves aient compris.

Enfin, l'internat permet d'exercer une forme de contrôle des élèves, qui se traduit par des différences d'aménagement de leur temps libre en dehors des cours. Les élèves à Sourdun consacrent chaque semaine deux heures de plus à leurs devoirs que les élèves témoins, passent cinq heures de plus en étude surveillée, sont deux fois plus nombreux à bénéficier d'un soutien individualisé de la part d'un de leurs professeurs, et passent trois fois moins de temps devant la télévision. On n'observe pas de phénomène de rattrapage le week-end, lorsque les élèves rentrent dans leurs familles. Au contraire, les habitudes prises pendant la semaine semblent se prolonger, puisque les élèves scolarisés à Sourdun consacrent aussi plus de temps à faire leurs devoirs le samedi que les élèves témoins.

Notre dispositif expérimental ne permet pas de dire quelle est la part que l'on peut attribuer à ces différentes composantes dans l'effet total produit par Sourdun. Cependant, il est vraisemblable que les effets très importants de l'internat sur l'emploi du temps et le cadre de

travail des élèves ont été bénéfiques. En outre, le fait d'être entouré de camarades motivés a probablement été positif aussi, même si le fait de ne plus être parmi les bons élèves de leur classe a pu dans un premier temps décourager certains élèves. Par ailleurs, notre dispositif expérimental nous permet aussi de montrer que les élèves scolarisés dans les établissements d'origine des internes n'ont pas été affectés par le départ de ces "bons éléments", sans doute parce qu'ils sont peu nombreux à avoir rejoint l'internat.

Une recherche qui en appelle d'autres : suivi de long terme et autres internes

Il semble crucial de prolonger notre travail par un suivi de long terme des élèves, afin de déterminer les effets de Sourduin sur leurs études supérieures et leur insertion professionnelle. De tels suivis sont désormais monnaie courante dans la littérature internationale, et ce sont souvent eux qui donnent les éléments les plus probants pour juger de l'efficacité d'une intervention.

Par ailleurs, notre étude concerne seulement l'internat d'excellence de Sourduin. Extrapoler ses impacts à d'autres internes reste hasardeux, les modes de fonctionnement variant fortement d'un établissement à l'autre. Si l'on souhaite porter un diagnostic plus large sur cette politique, il sera nécessaire de compléter notre travail par des études quantitatives portant sur d'autres internes d'excellence.

Un dispositif efficace, qui laisse ouverte la question des actions à mener à plus grande échelle.

Les résultats de l'internat d'excellence de Sourduin montrent qu'il est possible, par un dispositif ciblé, d'accroître significativement les compétences et l'ambition scolaires d'élèves d'origine modeste, battant ainsi en brèche l'image selon laquelle les politiques scolaires seraient impuissantes face aux inégalités.

Il n'en reste pas moins que cette politique fait aussi un choix : celui de concentrer des ressources importantes sur des élèves motivés et de niveau scolaire médian. Notre évaluation ne permet pas de dire si ce qui réussit avec eux réussirait aussi bien avec des élèves plus faibles ou moins motivés. Mais elle pose en creux la question des actions à mener auprès de ces derniers, dans un contexte où différents travaux montrent que, contrairement à ce qui se passe à Sourduin, les ressources supplémentaires par élève affectées à l'éducation prioritaire sont aujourd'hui très limitées.

Table des matières

1	Introduction	9
2	Une expérience contrôlée : méthodologie	12
2.1	Recrutement des élèves et tirage au sort	13
2.2	Scolarité effective à Sourdun et interprétation des effets	15
2.3	Significativité statistique et lecture des tableaux	17
2.4	Enquêtes et données	18
3	Le recrutement de Sourdun	24
4	La scolarité à l'internat d'excellence de Sourdun	29
4.1	L'emploi du temps et les conditions de travail à Sourdun	30
4.2	L'équipe éducative de Sourdun et ses relations avec les élèves	33
4.3	Les relations entre élèves à Sourdun	35
4.4	Sourdun : restaurer l'école?	37
5	Le coût supplémentaire d'une année de scolarité à Sourdun	37
6	L'impact de la scolarité à Sourdun sur les élèves bénéficiaires	41
6.1	Impact sur le niveau scolaire des élèves	41
6.1.1	Les tests standardisés et les scores construits par la DEPP	43
6.1.2	Au bout de deux ans, Sourdun produit un fort effet en mathématiques .	44
6.1.3	Ces effets sont cohérents avec les évaluations de dispositifs similaires . .	47
6.1.4	L'internat d'excellence de Sourdun comparé à des politiques alternatives	48
6.1.5	L'effet de Sourdun n'est pas significativement différent parmi les garçons et les filles, les collégiens et les lycéens, les élèves forts et les élèves faibles.	52
6.1.6	Impact sur les résultats au brevet.	53
6.2	Impact sur les aspirations, la motivation et la confiance en soi des élèves	54
7	L'impact de la scolarité à Sourdun sur les familles des élèves	58
7.1	Impact sur les parents	58
7.2	Impact sur la fratrie	60
8	L'impact de Sourdun sur les élèves qui n'en bénéficient pas	61
9	Conclusion	62

1 Introduction

La politique d'internat d'excellence constitue une orientation nouvelle et importante au sein des politiques scolaires. Elle s'inscrit en France dans une double histoire, celle de l'internat éducatif et celle de l'éducation prioritaire.

Après une longue phase de déclin (Robert, 2008, Glasman, 2010), le principe de l'internat a été réhabilité à partir des années 2000, notamment sous l'effet de la loi de programmation pour la cohésion sociale (2005), puis du plan *Espoir Banlieues* (2008). Ce plan a d'abord conduit à l'institution de "places labellisées", places réservées dans des internats existants aux jeunes issus des quartiers de la politique de la ville, puis à la création d'internats d'excellence spécifiquement dédiés. Alors que la formule de l'internat subsiste traditionnellement pour répondre à la demande d'élèves et de familles résidant loin d'un établissement qui offre une option ou une filière particulière, l'internat d'excellence vise quant à lui à développer un projet éducatif spécifique s'adressant à des élèves "pour lesquels la motivation de la demande d'internat n'est pas d'ordre géographique, mais éducative, sociale et/ou économique" (Robert, 2008). Sourdu est le premier internat dédié à avoir ouvert ses portes, à la rentrée 2009. Au total, 4 173 places étaient disponibles dans 45 internats d'excellence à la rentrée 2012¹.

L'internat d'excellence emprunte aussi à la tradition de l'éducation prioritaire. Il vise en effet à promouvoir l'égalité des chances pour des élèves d'origine modeste et encourage l'innovation pédagogique². Il se distingue cependant des formes initiales de cette politique par différents traits spécifiques, et notamment par un changement d'échelle et de ciblage. Là où des dispositifs comme les Zones d'éducation prioritaire ont mobilisé des moyens à peine supérieurs³ sur un vaste ensemble d'établissements dans le cadre d'une carte scolaire, l'internat d'excellence concentre, sur un petit nombre d'élèves issus des quartiers populaires, des moyens humains et financiers conséquents, dans le cadre d'un recrutement affranchi de la carte scolaire. En outre, cette politique vise des élèves motivés, avec un "potentiel" scolaire, mais dont on estime qu'ils ne bénéficient pas d'un environnement favorable à leur réussite scolaire, là où initialement l'accès à la politique d'éducation prioritaire ne reposait pas sur des critères de motivation ni de potentiel.

L'originalité de cette nouvelle "forme scolaire", ce "renouveau de l'éducation prioritaire" (Rayou et Glasman, 2012), et l'engagement budgétaire dont il a bénéficié, concentré sur un nombre relativement faible d'élèves⁴, justifie les efforts d'étude et d'évaluation qui com-

1. <http://www.education.gouv.fr/cid50541/les-internats-d-excellence.html>

2. Kherroubi et Rochex (2002) soulignent le rôle de l'innovation pédagogique au moment du lancement de la politique d'éducation prioritaire dans les années 1980.

3. Voir Bénabou, Kramarz, et Prost (2004), Befy et Davezies (2012), Cour des comptes (2010).

4. L'investissement sur les 13 premiers internats d'excellence est évalué à plus de 200 millions d'euros, pour environ un millier d'élèves en 2010-2011 (Delahaye et Dulot, 2011), à quoi s'ajoutent des coûts de fonctionne-

mentent à se déployer et auxquels ce rapport contribue. Un rapport d'inspection (Delahaye et Dulot, 2011) et, plus récemment, la vaste monographie dirigée par Rayou et Glasman (2012) dans le cadre de l'Institut français de l'éducation, permettent notamment de mieux connaître et comprendre le fonctionnement des internats d'excellence dans leur diversité, l'appropriation du dispositif par les personnels enseignants et éducatifs, l'attente des familles et l'usage qu'elles en font, la variété des trajectoires des élèves et les ajustements institutionnels que le dispositif a suscité. Ces travaux, et ceux qui viennent d'être lancés par le Fonds d'expérimentation pour la jeunesse, sont extrêmement importants, notamment parce qu'ils permettent de comprendre les décalages entre la politique telle qu'elle est définie en théorie et la réalité du terrain.

Ce rapport d'évaluation adopte une démarche différente pour répondre à d'autres questions. Il vise à mesurer formellement les impacts de deux années de scolarisation à l'internat d'excellence de Sourdun sur les élèves recrutés en 2009 et en 2010, au moment du lancement de l'internat. Il existe de multiples domaines d'impact possibles, mais on privilégie ici quatre d'entre eux : les performances scolaires des élèves ; leurs comportements et leurs aspirations ; les effets induits sur leurs familles ; enfin, d'éventuels impacts sur leurs camarades de classe restés dans leur établissement d'origine. Obtient-on de meilleurs résultats, acquière-t-on davantage de compétences cognitives lorsque l'on est scolarisé à Sourdun ? Développe-t-on une ambition scolaire spécifique ? Cela a-t-il des effets sur l'attitude des parents, sur la fratrie ? Les élèves scolarisés dans les établissements d'origine des internes partis à Sourdun sont-ils pénalisés par le départ de ces bons éléments ?

Impact s'entend ici au sens fort : nous comparons la situation des élèves scolarisés à Sourdun, un an puis deux ans après leur entrée à l'internat, avec la situation qu'ils auraient connue, aux mêmes moments, s'ils n'y avaient pas été scolarisés. Pour pouvoir faire cette comparaison, nous utilisons la méthode de l'assignation aléatoire. Le nombre de candidats à Sourdun ayant excédé le nombre de places disponibles, les élèves admis ont été tirés au sort⁵. Le tirage au sort crée ainsi deux groupes d'élèves, statistiquement identiques au moment de leur candidature, mais dont l'un, le groupe "test", va suivre une scolarité à Sourdun, tandis que le groupe dit "témoin" n'y sera pas scolarisé. Nous avons ensuite mis en oeuvre un ambitieux suivi des 395 jeunes ayant pris part à ce tirage au sort, comprenant notamment une interrogation annuelle en français et en mathématiques. Nous avons aussi interrogé chaque année les familles des élèves. En comparant les résultats des deux groupes au fil du temps, nous sommes en mesure de quantifier des différences qui peuvent rigoureusement s'interpréter

ment importants.

5. Un tel tirage au sort est une pratique courante aux Etats-Unis. Il est par exemple utilisé de façon systématique pour affecter les élèves dans les *charter Schools*, une catégorie d'écoles particulièrement demandées. À Sourdun, il n'a été pratiqué que pour les cohortes sur lesquelles porte cette évaluation.

comme un effet *causal* de la scolarisation à Sourdun, puisque rien ne distinguait les deux groupes initialement.

Il convient d'emblée de reconnaître trois limites de notre travail, qui sont autant de perspectives de recherches futures. Tout d'abord, nous n'avons pas pu inclure d'autres internats d'excellence dans notre étude. Certains internats accueillent trop peu d'élèves pour qu'une étude statistique les concernant puisse produire autre chose que des effets très imprécisément estimés, ce qui ne nous a pas semblé justifier les coûts d'une telle étude. D'autres internats n'ont pas connu l'excès de candidatures nécessaire pour mener à bien un tirage au sort. C'est par exemple le cas de l'internat de Cachan, qui devait initialement être inclus dans notre travail, mais où l'on n'a finalement observé qu'un léger excès de candidatures pour la classe de 4ème. Cette étude nous apprend donc quels sont les impacts d'une scolarité à l'internat d'excellence de Sourdun, mais ne nous dit rien sur les effets produits par les autres internats. Extrapoler les impacts de Sourdun semble d'autant plus hasardeux que les populations d'élèves visés et les modes de fonctionnement varient fortement d'un internat à l'autre (Rayou et Glasman, 2012). Si l'on souhaite porter un diagnostic plus large sur cette politique, il sera nécessaire de compléter notre travail par des études quantitatives portant sur d'autres internats d'excellence.

Ensuite, si ce rapport mesure l'impact total d'une scolarité à Sourdun, il ne permet pas de distinguer l'effet propre de chacun des éléments qui caractérisent cette scolarité. Comme on le verra plus loin, les élèves bénéficient à Sourdun de plusieurs éléments dont ils n'auraient pas (ou moins) bénéficié ailleurs : régime de l'internat, études surveillées, aide individuelle aux devoirs... En outre, les professeurs de Sourdun sont différents de ceux qui enseignent dans les établissements où sont scolarisés les élèves témoins. Ils sont notamment plus jeunes, plus diplômés, et ils ont été recrutés sur profil⁶. Enfin, Sourdun change profondément le groupe de pairs des élèves, qui se retrouvent entourés de camarades de même profil que le leur (élèves motivés, avec un "potentiel" scolaire et issus d'un milieu socialement défavorisé). Il nous est impossible de déterminer, dans l'effet total produit par Sourdun, la part due à chacune de ces composantes. Pour ce faire, il aurait aussi fallu tirer au sort, parmi les élèves admis à Sourdun, les élèves éligibles à de l'aide individualisée, ceux pouvant bénéficier d'études surveillées, etc., ce qui aurait évidemment été impossible en pratique. À tout le moins, nous avons apporté un soin particulier à décrire de façon précise en quoi la scolarité à Sourdun est spécifique. Si bien qu'à défaut de pouvoir identifier ce qui produit l' "effet Sourdun", nous pouvons mettre en face de cet effet les facteurs susceptibles de l'expliquer. Vouloir aller plus loin, en affirmant par exemple que c'est l'ambiance entre les élèves plutôt que les études surveillées qui explique

6. Le recrutement sur profil est parallèle au mouvement des enseignants, principalement basé sur des points d'ancienneté, et permet au proviseur de recruter des enseignants qui se sont portés candidats spécifiquement sur le projet de l'établissement. Il constitue une exception dans le système français.

l'effet de Sourdun, ce serait simplement spéculer, et faire dire à cette évaluation plus que ce qu'elle peut nous apprendre.

Enfin, ce rapport présente les impacts d'une scolarité à Sourdun sur les compétences et les aspirations des élèves au bout de deux ans. Il semble maintenant essentiel de compléter notre travail par un suivi de long terme des élèves, afin de déterminer l'impact de Sourdun sur leurs études supérieures et leur insertion professionnelle. De tels suivis de long terme sont désormais monnaie courante dans la littérature internationale, et ce sont souvent eux qui donnent les éléments les plus probants pour juger de l'efficacité d'une intervention.

L'organisation du rapport est la suivante. La section 2 précise la méthodologie utilisée dans cette étude : nous expliquons en détail le principe et les modalités du tirage au sort, ainsi que les enquêtes qui ont permis de suivre les jeunes pendant deux ans. Cette section vise à montrer que les effets présentés ensuite s'interprètent bien comme des effets causaux de la scolarité à Sourdun sur la population ciblée. Nous décrivons plus en détail ce ciblage dans la section 3 : l'origine sociale et le niveau scolaire initial des élèves retenus pour entrer à Sourdun sont décrits et comparés à différentes populations de référence (ensemble des élèves scolarisés sur le territoire, élèves scolarisés dans des établissements de l'éducation prioritaire, etc.). La section 4 documente les principales différences entre la scolarité à Sourdun et celle dont ont bénéficié les élèves témoins. Nous pouvons ainsi détailler les spécificités du dispositif évalué dans ses multiples composantes : emploi du temps, conditions de travail, ambiance de classe, relations avec les adultes,... Nous présentons ensuite le surcroît de dépense par élève à Sourdun, qui constitue la contrepartie de ces spécificités, notamment en raison du coût de l'internat lui-même. Puis nous présentons les impacts de la scolarité à Sourdun sur les compétences des élèves en mathématiques et en français, ainsi que sur un certain nombre d'attitudes et de comportements (section 6). Enfin, nous nous demandons si le départ d'un élève en internat a un effet sur ses parents et ses frères et sœurs (section 7), ainsi que sur ses camarades de classe restés dans son établissement d'origine (section 8).

2 Une expérience contrôlée : méthodologie

L'objectif de ce travail d'évaluation est de mesurer l'impact d'une scolarité à Sourdun sur les résultats scolaires et les comportements ou attitudes des élèves. Pour ce faire, il nous faut comparer la situation des internes une ou deux années après leur entrée en internat à celle d'élèves exactement semblables mais n'ayant pas été scolarisés à l'internat. Les différences observées entre les deux groupes peuvent alors s'interpréter comme l'effet *causal* de l'internat d'excellence, car elles ne peuvent pas provenir d'autres facteurs distinguant les populations.

Il serait incorrect de comparer les internes avec le reste des élèves de l'académie de Créteil,

ou même avec le reste des élèves issus des mêmes établissements, ou même encore avec des élèves ayant obtenu exactement les mêmes notes l'année précédente. En effet, la motivation et le désir de s'impliquer dans sa scolarité sont des facteurs extrêmement valorisés dans le processus de recrutement. Comparer un interne très motivé avec un élève qui avait les mêmes notes dans le même établissement, mais qui n'a pas eu la motivation pour rejoindre Sourdun, ce serait attribuer par erreur à Sourdun des différences entre ces deux élèves qui résultent en fait de cette différence de motivation.

Pour cette raison, l'évaluation proposée dans ce rapport s'appuie sur le tirage au sort des élèves admis à l'internat parmi un ensemble de candidats. Le tirage au sort crée deux groupes d'élèves initialement identiques dont l'un, le groupe "test", va suivre une scolarité à l'internat, tandis que l'autre, le groupe "témoin", n'y est pas admis. Il suffit alors de comparer la situation des élèves du groupe "test" à celle des élèves du groupe "témoin" pour déterminer quel a été l'effet de l'internat sur les élèves admis.

Dans cette section, nous décrivons d'abord le processus de recrutement des internes et la place du tirage au sort. Par rapport aux grands principes exposés ci-dessus, nous devons faire face à deux difficultés : d'une part, il existe une mobilité des élèves vers et hors Sourdun qui échappe partiellement aux orientations fixées par le tirage au sort. D'autre part, le tirage au sort crée deux groupes parfaitement similaires lorsque les effectifs sont très grands, mais le nombre de places à Sourdun est limité. Nous discutons donc des méthodes mises en oeuvre pour résoudre ces deux difficultés. Enfin, nous présentons le dispositif d'enquêtes et de collecte des données sur lequel s'appuie notre travail.

2.1 Recrutement des élèves et tirage au sort

Cette évaluation porte sur les deux premières cohortes recrutées à Sourdun, au printemps 2009 pour l'année scolaire 2009-2010, et au printemps 2010 pour l'année scolaire 2010-2011. Comme indiqué plus loin, chacune de ces cohortes est suivie pendant deux ans.

En 2009, Sourdun a recruté des élèves pour entrer en classes de 4e, 3e et 2de : 129 places étaient ouvertes. En 2010, le recrutement a porté sur des classes allant de la 6e à la 1ère : 150 nouvelles places étaient offertes et permettaient de terminer la montée en charge de l'établissement. À partir de 2011, le recrutement s'est limité au remplacement du flux sortant et nous n'avons pas suivi ces nouvelles cohortes. En 2009, 275 candidatures ont été reçues ; en 2010, la notoriété de l'établissement s'étant accrue, 499 candidats ont adressé leurs dossiers. Les candidats devaient nécessairement provenir d'établissements de l'académie de Créteil.

Tous les ans, une commission de recrutement s'est réunie pour étudier les dossiers et écarter ceux qui n'étaient pas conformes au profil visé par la politique, c'est-à-dire celui d'élèves

motivés⁷, mais dont la situation sociale ou familiale permettait de penser que l'internat leur offrirait des conditions de travail plus propices au développement de leur potentiel. En outre, une lettre de motivation permettait d'évaluer l'engagement des jeunes et de leurs familles. 73 dossiers sur 275 ont été écartés en 2009, et 216 sur 499 en 2010, parce qu'ils n'étaient pas conformes aux attentes. Au total, ce sont donc un peu plus de 37% des dossiers qui ont été déclarés irrecevables par ces commissions. Quelques dossiers (5 en 2009 et 7 en 2010) ont par ailleurs été jugés prioritaires.

À l'issue de cette première étape, le nombre de candidats restant excédait le nombre de places encore ouvertes. Ceux à qui l'internat a été proposé ont alors été tirés au sort. Les autres ont été placés sur une liste d'attente, dans un ordre défini aléatoirement. Afin de respecter les contraintes d'organisation de Sourduin, liées notamment à la composition des chambrées et aux places disponibles dans chaque classe, le tirage au sort a été réalisé par strate (garçons ou filles, classe demandée). En cas de désistement, les candidats étaient appelés suivant l'ordre de la liste d'attente de leur strate. Notre groupe "test", que l'on qualifiera dans ce qui suit de "groupe orienté vers l'internat", se compose des élèves à qui une place en internat a été proposée, soit directement, soit à la suite d'un désistement. Notre groupe témoin se compose d'élèves classés en queue de liste d'attente, auxquels on n'a jamais proposé de place en internat. Le fait que l'ordre sur les listes soit aléatoire nous assure que les deux groupes ainsi constitués sont bien comparables.

En 2010, un deuxième internat d'excellence a ouvert dans l'académie de Créteil, à Cachan. Le recrutement a porté sur les classes de 4e et de 3e. Le projet initial était de compléter notre échantillon avec une cohorte de candidats à Cachan, mais pour cet internat, seule la classe de 4e a fait l'objet d'un léger excès de candidatures. Faute de disposer d'un échantillon suffisant pour estimer de façon précise l'effet de Cachan, nous avons jugé qu'il n'était pas pertinent d'inclure ces élèves dans notre analyse.

L'échantillon final de l'évaluation est donc constitué de 395 élèves candidats à Sourduin et retenus par les commissions. Ils se répartissent comme suit : 179 candidats appartiennent à la première cohorte recrutée en 2009 ; 216 candidats font partie de la deuxième cohorte recrutée en 2010. Le groupe orienté vers l'internat compte 258 élèves et le groupe témoin 137.

Le tableau 1 compare les deux groupes sur une trentaine de variables. Il s'agit de vérifier que le tirage au sort a bien créé deux groupes d'élèves statistiquement similaires⁸. Dans la première colonne figure la valeur moyenne de chacune de ces variables parmi les élèves témoins, tandis que la deuxième colonne donne la différence entre la moyenne du groupe orienté vers l'internat et celle du groupe témoin. La troisième colonne reporte l'écart-type de cette

7. Le niveau scolaire des candidats sera documenté en détail plus loin.

8. Cette table pourrait aussi permettre de décrire la population d'intérêt, mais nous revenons plus précisément sur cette question dans la section 3.

différence et la dernière colonne le nombre d'observations utilisées. Naturellement, les différences ne sont jamais exactement nulles, ce qui signifie qu'avec une taille d'échantillon donnée (395 élèves au maximum ici), les groupes ne sont jamais *parfaitement* semblables. Lorsque nous jugerons des impacts de Sourdu, nous devons tenir compte de ce bruit statistique comme cela est expliqué plus bas. Mais les différences observées ne sont pas plus grandes que celles que l'on obtiendrait si l'on tirait au sort deux échantillons dans deux groupes parfaitement semblables. Certaines différences (4) sont "significatives" et sont alors accompagnées d'une étoile, mais leur nombre est également conforme à ce que l'on attend du fait des aléas d'échantillonnage.

2.2 Scolarité effective à Sourdu et interprétation des effets

Une partie des élèves orientés vers l'internat n'ont pas été scolarisés à Sourdu, soit parce qu'ils se sont désistés, soit parce qu'ils ont quitté l'établissement rapidement, de leur propre fait ou à la suite d'une mesure disciplinaire. En outre, certains ont pu passer une année plutôt que deux dans l'établissement. Très peu d'élèves du groupe témoin sont finalement entrés à Sourdu; les quelques exceptions correspondent principalement à des rapprochements de fratrie. Le tableau 2 présente les taux de scolarisation effective à Sourdu dans les deux groupes. Un an après le tirage au sort (c'est-à-dire en mai 2010 pour la première cohorte et en mai 2011 pour la seconde), 77% des élèves orientés vers l'internat étaient effectivement scolarisés à Sourdu. Deux ans après, ce taux est de 63%. Dans le groupe témoin, le taux de scolarisation à l'internat est de 4% après un an et de 9% après deux ans.

Pour mesurer les effets de la scolarité à Sourdu, nous allons donc utiliser une technique classique lorsque le tirage au sort n'engendre pas une obligation d'entrer dans le dispositif, ce qui est fréquemment le cas. Nous avons deux groupes comparables de 258 et 137 jeunes respectivement. Du fait de l'intervention, ces deux groupes deviennent différents en ce que la grande majorité du groupe orienté vers l'internat est scolarisée à Sourdu tandis que c'est le cas d'un très petit nombre d'élèves dans le groupe témoin. Si la scolarité à Sourdu n'avait aucun impact, on ne devrait observer aucune différence significative entre ces deux groupes d'élèves un an ou deux ans après le tirage au sort. Inversement, si cette scolarité a des impacts positifs, cet effet se retrouvera dans la comparaison entre les deux groupes; il sera néanmoins dilué dans la mesure où seulement une partie des élèves orientés vers Sourdu y sont effectivement scolarisés.

Par conséquent, nous ne présentons pas directement les différences moyennes entre les deux groupes, mais nous les normalisons par la différence entre les proportions d'élèves effectivement scolarisés à Sourdu (par exemple 77%-4% au bout d'un an). Les coefficients que nous commentons s'interprètent alors comme l'effet moyen d'être scolarisé à Sourdu, et non

TABLE 1 – Comparaison des deux groupes formés par le tirage au sort.

	Moyenne Témoin	Orientés - Témoins	σ	N
Résultats scolaires et options l'année du tirage				
Note de français	12.209	-0.125	0.343	392
Ecart à la moyenne de classe en français	2.123	-0.090	0.328	373
Note de mathématiques	12.437	0.189	0.393	392
Ecart à la moyenne de classe en mathématiques	2.266	0.264	0.438	373
Latiniste ou helléniste	0.286	-0.097*	0.059	375
Germaniste	0.266	-0.051	0.056	375
Résultats au brevet l'année du tirage ou avant				
Mention TB au brevet	0.139	-0.065	0.048	211
Mention B au brevet	0.258	0.001	0.076	211
Mention AB au brevet	0.273	0.072	0.068	211
Sans mention au brevet	0.301	-0.038	0.077	211
Echec au brevet	0.000	0.026*	0.013	211
Note à l'épreuve terminale de français	11.749	-0.870*	0.462	149
Note à l'épreuve terminale de mathématiques	9.612	-0.301	0.589	149
Note à l'épreuve terminale d'histoire-géographie	11.392	-0.734	0.584	149
Indicateurs de comportement l'année du tirage				
Note de comportement	15.304	0.390	0.508	343
Ecart à la moyenne de classe en comportement	0.238	0.683	0.552	325
Nombre d'absences par trimestre	5.596	0.600	0.687	350
Origine sociale et environnement familial				
Boursier	0.404	0.013	0.062	393
Demi-pensionnaire	0.503	0.010	0.343	369
Nombre d'enfants du responsable légal	2.827	-0.100	0.196	393
Parents divorcés (EP ^a)	0.264	-0.005	0.063	351
Responsable légal sans diplôme (EP)	0.091	-0.002	0.050	343
Responsable légal a le bac ou plus (EP)	0.253	0.025	0.061	343
Famille monoparentale (EP)	0.372	-0.049	0.065	352
Seul le français est parlé en famille (EP)	0.386	0.090	0.064	352
Décision de candidater à l'internat (EP)				
Idée du collège/d'un enseignant	0.483	-0.116*	0.065	352
Idée de l'enfant	0.220	0.088	0.054	352
Idée des parents	0.238	-0.024	0.046	352

a. EP signifie que la variable est issue de l'enquête parents.

TABLE 2 – Effectifs à Sourdu

	1 an après le tirage	2 ans après le tirage
Élèves orientés vers l'internat	258	258
Élèves orientés scolarisés à Sourdu	199 (77.13%)	163 (63.18%)
Élèves témoins	137	137
Élèves témoins scolarisés à Sourdu	6 (4.38%)	12 (8.76%)
Nombre total d'élèves	395	395
Élèves scolarisés à Sourdu	205 (51.90%)	175 (44.30%)

l'effet d'avoir 77% de chances plutôt que 4% d'y être scolarisé. Ces effets sont mesurés au sein du sous-groupe d'élèves qui s'est conformé au tirage au sort, c'est-à-dire ceux qui, étant aléatoirement orientés vers l'internat, y ont effectivement été scolarisés⁹. Notre rapport ne nous apprend donc pas quels auraient été les effets d'une scolarité à l'internat d'excellence de Sourdu sur les élèves qui ne se sont pas conformés au tirage, par exemple ceux qui se sont désistés. Cela ne nous semble pas être une limite majeure, puisque cela consiste à évaluer le dispositif tel qu'il est, et non un dispositif hypothétique qui forcerait les jeunes initialement candidats à rester en internat d'excellence.

2.3 Significativité statistique et lecture des tableaux

Comme toutes les études de ce genre, l'exercice auquel nous nous livrons est un exercice statistique. Le tirage au sort annulerait toutes les différences entre le groupe d'élèves orientés vers l'internat et le groupe témoin si ces deux groupes étaient composés d'une infinité d'élèves. Ce n'est évidemment pas le cas, et de petites différences dues au hasard peuvent subsister entre ces deux groupes. C'est par exemple ce que l'on observe dans la table 1. Lorsque nous comparerons les deux groupes un ou deux ans après le tirage au sort, nous nous attacherons à mettre en évidence les différences statistiquement "significatives" aux seuils usuels de 5 ou 10%, c'est-à-dire celles qui ont de faibles chances (moins de 5 ou 10%) d'être dues au hasard.

Étant donnée la petite taille de notre échantillon, le seul aléa du tirage au sort peut faire apparaître des différences relativement importantes entre le groupe d'élèves orientés vers l'internat et le groupe témoin, si bien que nous ne pouvons considérer comme significatives que des différences vraiment conséquentes entre ces deux groupes. Notre dispositif expérimental permettra donc de détecter les effets de Sourdu s'ils sont forts, mais pas s'ils sont modestes,

9. Dans la littérature statistique, cette méthode est appelée "estimation à variables instrumentales" (voir Imbens et Angrist, 1994).

puisque nous ne serons pas capable de distinguer de faibles effets d'un bruit statistique¹⁰. C'est une limite de notre protocole, mais au lancement de ce projet on a considéré que l'investissement dans l'internat d'excellence de Sourdun était suffisamment important pour justifier un protocole expérimental capable de détecter seulement de forts effets, puisque seuls des effets importants semblaient en mesure de justifier ces investissements.

La grande majorité de nos tableaux sont présentés sur le modèle suivant. La première colonne liste les variables considérées dans le tableau. La seconde colonne, intitulée "Moyenne Témoin", contient la moyenne de ces variables pour les élèves du groupe témoin. La troisième colonne, intitulée "Sourdun - Témoin" présente l'effet de Sourdun sur ces différentes variables, estimé à partir de la méthode décrite ci-dessus. Un coefficient positif signifie que la scolarité à Sourdun a un effet positif sur la variable considérée, tandis qu'un coefficient négatif signifie le contraire. Pour juger de l'importance d'un effet, il sera souvent utile de le comparer au chiffre de la première colonne, qui s'interprète comme la situation qu'auraient connue les élèves s'ils n'avaient pas été scolarisés à Sourdun. Un effet de +5 n'a pas la même interprétation si la situation de référence est de 5 ou si elle est de 100. Dans le premier cas, l'effet correspond à une augmentation de 100%, contre seulement 5% dans le second cas.

Dans cette seconde colonne, nous incluons aussi le résultat d'un test statistique de nullité de cet effet. Le nombre d'étoiles placées à côté du coefficient indique son niveau de significativité statistique. L'absence d'étoile signifie que l'on ne peut pas rejeter l'hypothèse d'absence d'effet. S'il existe une différence, celle-ci est trop faible pour être détectée et on ne peut pas la distinguer d'un simple bruit statistique. Une étoile indique qu'il y a moins de 10% de chances que la différence entre les deux groupes soit due à de l'aléa d'échantillonnage. Deux ou trois étoiles indiquent que cette probabilité est inférieure ou égale à 5% et à 1% respectivement.

La troisième colonne, intitulée σ , présente l'écart-type du coefficient présenté dans la seconde colonne. C'est de lui que dépend la significativité statistique de ce coefficient. Enfin, la dernière colonne reporte le nombre d'observations utilisées pour estimer l'effet présenté dans la deuxième colonne. Pour les variables qui concernent les élèves de notre échantillon, ce chiffre est en général un peu inférieur à 395 puisqu'il y a souvent quelques élèves pour lesquels l'information est manquante. La figure 1 propose une aide à la lecture des tableaux du rapport, en prenant pour exemple la table 20.

2.4 Enquêtes et données

Une caractéristique importante de ce travail est que nous avons suivi pendant deux années notre échantillon, en interrogeant les élèves et leurs familles sur un grand nombre de

10. Plus précisément, avec l'échantillon de 395 élèves dont nous disposons, nous avons 80% de chances de détecter une différence de 35% d'un écart-type entre les deux groupes, au seuil de significativité usuel de 95%.

TABLE 20 - Études supérieures : ambition et confiance en soi

	Moyenne Témoin	Sourdun - Témoin	σ	N
Ambition : l'élève souhaite...				
Faire des études supérieures	0.800	0.038	0.060	362
Intégrer une classe préparatoire	0.097	0.177***	0.064	362
Obtenir au moins un master	0.571	0.143*	0.082	362
Confiance en soi : l'élève se sent capable...				
D'obtenir le baccalauréat	0.886	0.004	0.047	361
D'intégrer une classe préparatoire	0.409	0.055	0.092	360
D'obtenir un master	0.518	0.145	0.090	360

Aide à la lecture :

1ère colonne : **9.7%** des élèves du groupe témoin souhaitent intégrer une classe préparatoire.

2ème colonne : les élèves scolarisés à Sourdun sont **17.7 points de pourcentage plus nombreux** à avoir cette ambition. Cela signifie que 27.4% d'entre eux veulent intégrer une classe préparatoire (9.7%+17.7%). Les trois étoiles indiquent que la différence entre les moyennes des deux groupes d'élèves est statistiquement significative au seuil de 1%.

3ème colonne : l'écart-type de la différence de moyennes entre les deux groupes est de **6.4** points de pourcentage.

4ème colonne : l'analyse a été réalisée à partir des réponses de **362** élèves.

FIGURE 1 – Aide à la lecture des tables

dimensions. Nous présentons maintenant ces enquêtes¹¹.

Données décrivant la situation initiale des élèves

Nous disposons de données individuelles nous permettant de caractériser l'environnement scolaire et familial des élèves avant leur entrée dans l'expérimentation. Les données scolaires (moyenne de contrôle continu, note de comportement et assiduité, langues et options...) ont été recueillies pour l'essentiel à partir des bulletins de note transmis dans les dossiers de

11. Le traitement des données a fait l'objet d'une déclaration à la CNIL. Aucune des bases sur lesquelles nous travaillons ne contient de données nominatives et, naturellement, aucune des tables présentées ici ne permet d'identifier des élèves ou des groupes d'élèves. Les données d'origine sont conservées sous un format crypté.

candidature à l'internat. Le rectorat de Créteil nous a également communiqué les résultats au brevet des élèves. Ces résultats comprennent les décisions du jury d'admission (échec, admis sans mention...) ainsi que les notes des élèves au contrôle continu et aux épreuves terminales, à condition que leurs responsables légaux aient autorisé la divulgation de ces notes.

Nous disposons également d'un certain nombre de caractéristiques sociodémographiques des élèves candidats à l'internat (PCS et nombre d'enfants du responsable légal, nombre de parts de bourse...). Elles nous ont été transmises par le rectorat de Créteil à partir de l'exploitation du fichier Sconet.

Tests standardisés et questionnaires auprès des élèves

Les élèves de notre échantillon ont été soumis deux fois à une batterie de questionnaires, un an puis deux ans après le tirage au sort. Il s'agissait d'évaluer leurs compétences en français et en mathématiques, de mesurer un certain nombre de comportements et attitudes (motivation, désir de faire des études supérieures, estime de soi...), et de décrire leurs conditions de scolarisation. Les élèves ayant candidaté au printemps 2009 ont donc été interrogés en mai 2010 puis en mai 2011, tandis que ceux qui ont candidaté au printemps 2010 ont passé ces tests en mai 2011 et en mai 2012.

Chaque année, les questionnaires étaient passés au cours de deux sessions d'une heure et trente minutes, surveillées par des assistants d'éducation. La première session était consacrée à un test de français d'une durée d'une heure, suivi d'un questionnaire non cognitif pendant la demi-heure restante. Lors de la seconde session, les élèves passaient un test de mathématiques pendant une heure, avant de répondre à un second questionnaire. La grande majorité des élèves a passé ces questionnaires sur ordinateur¹².

Les tests étaient passés dans l'établissement scolaire de l'élève. Organiser la passation a été relativement aisé pour les élèves du groupe orienté vers Sourdu, puisque la majorité d'entre eux se trouvaient sur place. En revanche, les élèves du groupe témoin étaient dispersés dans un grand nombre d'établissements, et l'organisation des tests a donc été beaucoup plus compliquée pour eux. Au total, des tests ont été passés dans 55 établissements scolaires en 2010, 137 en 2011, et 102 en 2012, si bien que pas moins de 169 établissements différents ont été concernés par une passation au moins une fois au cours de ces trois années. Ils sont pour la plupart regroupés dans l'académie de Créteil, comme le montre la figure 2, mais certains se situent en dehors car des élèves ont changé d'établissement au cours des deux années pendant lesquelles nous les avons suivis, par exemple à la suite d'un déménagement.

Avant chaque passation, l'équipe d'évaluation a contacté ces établissements par téléphone afin d'expliquer comment faire passer les tests (système informatique requis, matériel autorisé

12. Certains tests ont du être passés sur papier à cause de problèmes informatiques.

pour les élèves, surveillance...). Il s'agissait aussi de tester le bon fonctionnement du site Internet sur lequel étaient hébergés les tests : parfois l'accès à notre site était interdit par un pare-feu de l'établissement qu'il fallait alors désactiver. Ces appels servaient enfin à tirer au sort trois camarades de classe de l'élève de l'échantillon concerné par nos tests, afin qu'ils les passent avec lui. L'objectif était d'éviter de stigmatiser l'élève concerné en lui faisant passer le test tout seul, et de décrire le niveau scolaire des camarades de classe des élèves témoins. L'équipe de recherche a aussi envoyé des assistants dans les établissements qui estimaient avoir besoin d'aide pour faire passer les tests. L'organisation des tests a donc nécessité la mise en place d'une logistique importante.

Pour garantir la validité de l'évaluation, il était très important de réussir à faire passer les tests à la quasi-totalité des élèves de l'échantillon. En effet, le pourcentage d'élèves interrogés allait par construction être proche de 100% parmi les élèves orientés vers l'internat : la grande majorité d'entre eux s'y trouvait, il était donc très facile de les localiser pour les interroger. Notre analyse aurait pu être fortement biaisée si nous avions seulement réussi à interroger 60 ou 70% des élèves du groupe témoin. Nous aurions en effet comparé à la quasi-intégralité du groupe d'élèves orientés vers l'internat un groupe témoin amputé des élèves les plus difficiles à localiser et les plus absentéistes. Cela nous aurait sans doute amené à sous-estimer l'effet de l'internat, puisque l'on aurait comparé les meilleurs élèves du groupe témoin à l'ensemble du groupe d'élèves orientés vers l'internat.

Comme le montre la table 3, plus de 90% des élèves de notre échantillon ont passé nos tests un an après leur candidature, et un peu moins de 90% les ont passés deux ans après. Il convient de souligner que cet excellent résultat a été obtenu grâce à la coopération exceptionnelle de l'ensemble des établissements scolaires impliqués, ainsi qu'au soutien indéfectible du rectorat de Créteil et de la DGESCO. Cette table montre aussi que les taux de passation ne sont pas significativement différents parmi les élèves orientés à l'internat et parmi ceux du groupe témoin, ce qui limite le risque de biais évoqué au paragraphe précédent.

Cette enquête est donc très exceptionnelle par son ambition, en raison de l'objectif pratiquement atteint de suivre tous les élèves où qu'ils soient. Ceci s'est traduit par un coût important, reflétant notamment le recrutement de nombreux assistants de recherche pour organiser les passations dans les établissements, le recours à un prestataire d'enquête et une logistique importante. Le projet a coûté 154 000 euros par an pendant trois ans et demi. Toutefois, ceci représente moins de 0.5% du coût *additionnel* de Sourdun par rapport à une scolarité normale (à peu près 10 000 euros supplémentaires \times 429 élèves, soit 4.3 millions d'euros par an), si on calcule ce coût en valeur actualisée avec un taux d'actualisation de 5%.

TABLE 3 – Taux de réponse au dispositif d'enquêtes

	Moyenne Témoin	Orientés - Témoins	σ	N
1 an après le tirage				
Test de français	0.928	-0.024	0.019	395
Questionnaire 1	0.916	-0.021	0.021	395
Test de mathématiques	0.922	-0.029	0.020	395
Questionnaire 2	0.916	-0.028	0.021	395
Enquête parents	0.735	0.100*	0.054	395
2 ans après le tirage				
Test de français	0.904	-0.021	0.022	395
Questionnaire personnel 1	0.892	-0.019	0.024	395
Test de mathématiques	0.887	-0.013	0.029	395
Questionnaire personnel 2	0.887	-0.016	0.029	395
Enquête parents	0.699	0.026	0.050	395

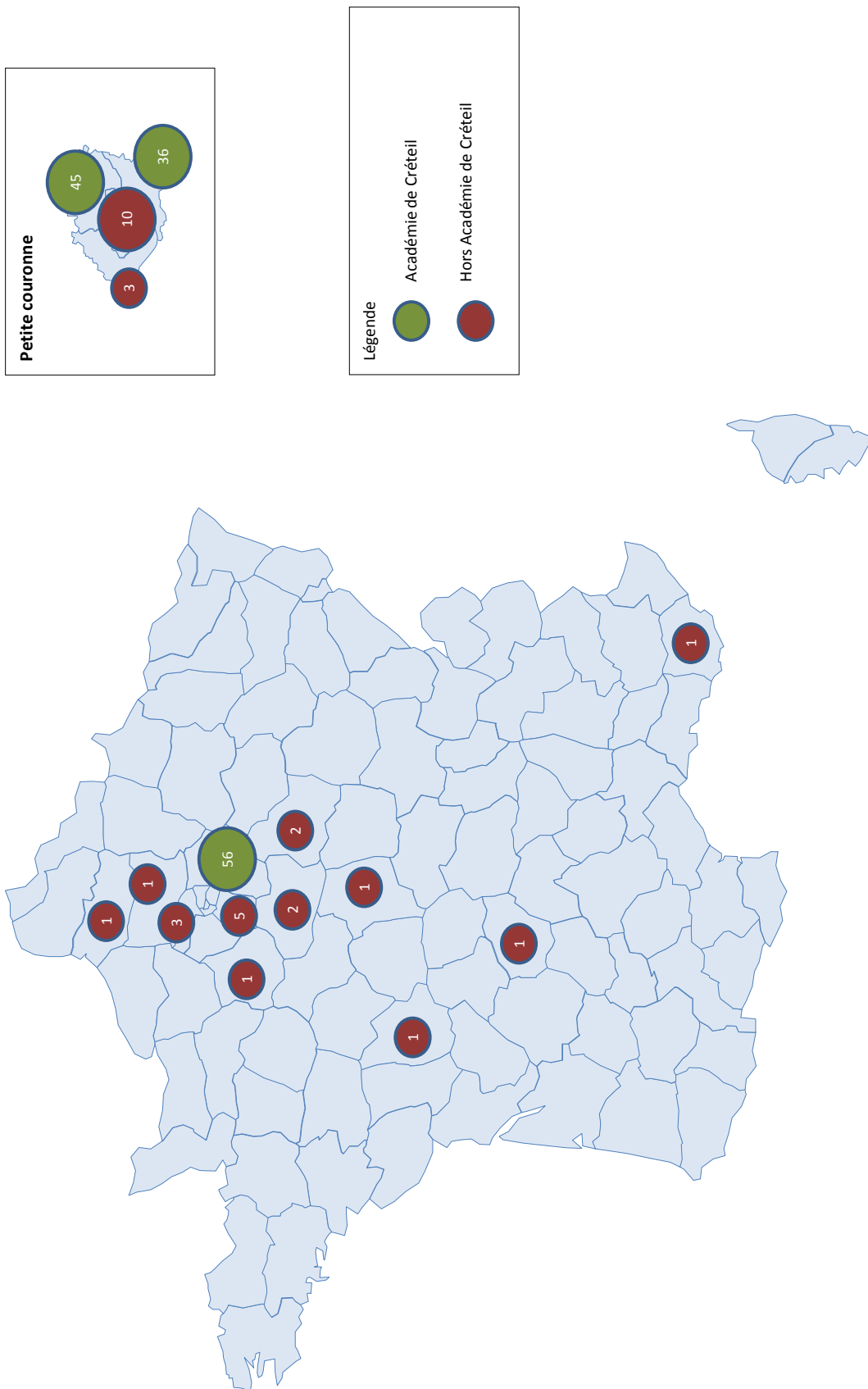


FIGURE 2 – Carte des 169 établissements où des tests ont été passés.

L'enquête auprès des familles

Un questionnaire d'une durée moyenne de 30 minutes, administré par téléphone, a également été soumis par un prestataire aux parents des élèves de la population expérimentale. Une première série de questions portait sur l'enfant candidat à l'internat, afin d'identifier qui était à l'origine de sa candidature et quelles en étaient les motivations. L'enquête visait également à mesurer l'opinion des parents sur la scolarité et le comportement de leurs enfants, ainsi que leur désir de les voir faire des études supérieures. Une troisième série de questions servait à caractériser le niveau de formation des parents de l'élève, ainsi que leur situation professionnelle. Comme pour les tests soumis aux élèves, cette enquête s'est déroulée une première fois un an après le tirage au sort, une seconde fois deux ans après.

Un an après le tirage au sort, le taux de réponse des parents dont l'enfant a été orienté vers l'internat est supérieur de 10 points de pourcentage à celui des parents d'élèves du groupe témoin (voir table 3). Cela reflète sans doute le fait que les parents d'élèves scolarisés à Sourdu se sont sentis plus concernés par notre enquête. Cette différence étant statistiquement significative, il est possible que les parents d'élèves témoins qui ont répondu à l'enquête ne soient pas comparables aux parents d'élèves du groupe orienté vers l'internat qui y ont répondu, si bien que la comparaison des réponses des parents des deux groupes pourrait capturer autre chose que le seul effet de Sourdu. Nous n'avons donc pas mobilisé les données de cette première enquête dans notre analyse. Deux ans après le tirage au sort, le taux de réponse moyen est un peu plus faible, de l'ordre de 70% environ, mais la différence de taux de réponse s'est considérablement réduite : le taux de réponse des parents du groupe d'élèves orientés vers l'internat est similaire à celui des parents d'élèves du groupe témoin. Ce sont donc les données de cette seconde vague d'enquête que nous mobilisons dans la suite de ce rapport, et notamment dans la section 7 qui considère l'impact de Sourdu sur les familles des élèves.

3 Le recrutement de Sourdu

Le ciblage des bénéficiaires est une particularité forte des internats d'excellence, qui les distingue de la tradition de l'éducation prioritaire. L'internat d'excellence vise les élèves motivés et désireux d'apprendre, mais qui ne disposent pas, dans leur quotidien, des conditions nécessaires à leur épanouissement scolaire et personnel. Le niveau scolaire attendu n'est pas explicitement défini par l'Éducation nationale, mais une ambiguïté a rapidement été levée : l'internat n'est pas destiné à accueillir les élèves excellents, mais à viser l'excellence pour les élèves qu'il accueille (Delahaye et Dulot, 2011).

Au total, aucune des conditions de recrutement n'est formellement définie, et la notion d'élèves "méritants", parfois mentionnée, est loin d'être simple à établir (Rayou et Glasman,

2012). Le profil des jeunes qui sont finalement entrés à l'internat de Sourdun résulte de plusieurs niveaux de sélection successifs. D'abord, les établissements d'origine ont invité les familles à candidater, mais les familles elles-mêmes ont pu en prendre l'initiative : la moitié des familles nous ont déclaré que l'idée de l'internat venait de l'établissement et l'autre moitié qu'elle venait des parents ou de l'enfant (cf. table 5). Ensuite, les commissions de recrutement (composées de représentants de l'autorité académique, du chef d'établissement et d'une assistante sociale) ont retenu les dossiers sur la base de la situation familiale, de la motivation de l'élève (attestée par une lettre jointe au dossier) et de son bulletin de notes. Si l'on est en droit de penser que les établissements prescripteurs ont orienté vers l'internat des élèves correspondant au profil défini par la politique, aux ambiguïtés près de sa définition, il est possible que les candidatures "spontanées" aient émané d'élèves très différents de ceux visés. Les commissions de sélection qui se sont tenues chaque année ont toutefois pu servir de filtre, puisqu'elles ont écarté un nombre important de candidats considérés comme "hors profil". Le rapport de Rayou et Glasman (2012) contient une intéressante monographie d'une commission de recrutement qui révèle les difficultés de l'exercice pour les acteurs et les incertitudes du résultat. Il faut noter que la difficulté a pu être exacerbée à l'internat de Sourdun, l'un des seuls à avoir eu à gérer un excès sensible de candidatures. Nos données nous permettent de définir précisément –en termes d'appartenance sociale et de niveau scolaire initial– la population qui a résulté de ce processus durant les deux premières années de recrutement à Sourdun.

La table 4 compare l'origine sociale des candidats retenus¹³ avec ceux de différentes populations de référence : l'ensemble des élèves sur le territoire national, les élèves scolarisés dans des établissements de l'éducation prioritaire, ceux de l'académie de Créteil, et leurs camarades de classe d'origine l'année où ils ont candidaté. Il ressort tout d'abord que les élèves retenus sont plutôt des élèves socialement défavorisés : la proportion d'élèves boursiers parmi les collégiens de notre échantillon est presque deux fois plus élevée que parmi l'ensemble des collégiens, et proche de celle que l'on observe dans les établissements de l'éducation prioritaire. Pour les lycéens, le constat reste sensiblement le même (cf. table 4, premier panel). Cette proportion est légèrement inférieure à celle que l'on trouve dans les 13 internats d'excellence existant en 2011 d'après Delahaye et Dulot (2011) et comparable à celle que l'on observe dans les établissements de l'éducation prioritaire¹⁴.

13. Les élèves retenus sont bien ici tous ceux qui, ayant candidaté, ont été retenus par la commission. Les statistiques présentées portent sur l'ensemble de ces élèves, qu'ils aient été orientés vers Sourdun ou non par le tirage au sort.

14. Le taux d'élèves boursiers parmi les candidats retenus peut sembler faible, puisque les internats d'excellence sont censés s'adresser en priorité à des élèves d'origine modeste. Il convient cependant de garder à l'esprit que certaines familles éligibles ne déposent sans doute pas de demande de bourse, en raison de la complexité des procédures, même s'il n'y a pas de statistiques disponibles au niveau national documentant l'ampleur de ce phénomène. Aussi, la proportion de boursiers reflète-t-elle peut-être imparfaitement l'origine sociale des élèves retenus par la commission de sélection.

En ce qui concerne la catégorie socio-professionnelle de leurs parents, la proportion d'élèves dont les responsables légaux sont sans activité, une catégorie qui inclut notamment les chômeurs, est plus élevée de 10 points de pourcentage parmi les élèves retenus que parmi leurs camarades de classe d'origine, ou que dans l'ensemble de l'académie de Créteil (cf. table 4, second panel). La proportion d'enfants d'ouvriers et d'employés est sensiblement identique au reste de l'académie. La table 5 présente d'autres caractéristiques qui viennent compléter la description de l'origine sociale des candidats retenus : 36.9% vivent dans des familles monoparentales, leurs responsables légaux ont en moyenne un peu moins de 3 enfants à charge, et 28.8% d'entre eux ont un diplôme supérieur ou égal au baccalauréat. Le français est la seule langue parlée en famille pour moins de la moitié des candidats.

En termes de niveau scolaire initial, les candidats retenus étaient sensiblement meilleurs que leurs camarades de classe. L'année de leur candidature, ils se situaient 2 points au-dessus de la moyenne de leur classe en français, et 2.5 points au-dessus en mathématiques, avec des moyennes de 12.5 et 13 respectivement. Pour ceux qui ont passé le diplôme national du brevet (DNB), nous pouvons également utiliser leurs notes au contrôle continu ainsi que celles de leurs camarades de classe pour déterminer le rang auquel ils se situaient dans leur classe. Nous avons ainsi calculé la moyenne de leurs rangs, et avons obtenu les valeurs reportées dans le tableau, de 0.293 en français et de 0.314 en mathématiques¹⁵. Cela signifie que les élèves retenus se situaient à peu près au troisième décile de leur classe, soit entre le premier quart et le premier tiers. En d'autres termes, les candidats étaient en moyenne 7ème ou 8ème de leur classe sur à peu près 25 élèves (cf. table 4, troisième panel).

Le graphique 3 présente la distribution des notes de contrôle continu en français et en mathématiques des candidats retenus. En français, elle est très resserrée. Seuls 4% des candidats ont une note inférieure à 8 sur 20, et moins de 16% ont moins de 10. Les deux tiers ont entre 10 et 15, et seuls 19% d'entre eux ont plus de 15. En mathématiques, les notes sont un peu moins resserrées. Seuls 19% d'entre eux ont moins de 10, mais 31% ont plus de 15. L'internat d'excellence accueille donc une population relativement homogène, composée en majorité d'assez bons élèves dans leur classe d'origine.

Les candidats retenus étaient aussi légèrement au-dessus de la moyenne de l'ensemble des élèves scolarisés en France. Parmi ceux qui avaient passé le DNB avant de candidater à Sourdun, seuls 1.7% avaient échoué, contre 19.0% dans l'ensemble de la population (table 4, troisième panel). De même, la proportion de mentions Bien ou Très Bien au brevet parmi les élèves retenus est supérieure à celle que l'on observe dans l'ensemble du pays. Enfin, leurs notes aux épreuves terminales du DNB sont approximativement 14 points d'écart-type au-dessus de la moyenne française, et 40 points d'écart-type au-dessus de celle de leurs camarades

15. Notre variable vaut 0 si l'élève est premier de sa classe et 1 s'il est dernier.

TABLE 4 – Origine sociale et niveau scolaire des candidats retenus

	Candidats Retenus	France ^a	Éducation prioritaire	Créteil	Camarades
Taux de boursiers					
Collégiens	0.435	0.278	0.468	0.272	
Lycéens	0.413	0.249			
PCS du responsable légal ^b					
Employés	0.257			0.211	0.210
Ouvriers	0.270			0.258	0.278
Sans activité	0.178			0.071	0.082
Contrôle continu ^c					
Français	12.566				10.500
Rang en français	0.293				
Mathématiques	13.002				10.529
Rang en mathématiques	0.314				
Brevet des collèges					
<i>Décision</i>					
Echec (si présent)	0.019	0.190	0.262	0.230	0.232
Reçu, mentions B et TB (si admis)	0.369	0.267	0.199	0.223	0.220
<i>Notes aux épreuves terminales</i>					
Français	0.135	0.000	-0.288	-0.182	-0.335
Mathématiques	0.145	0.000	-0.352	-0.253	-0.241

a. Sauf mention contraire, les données concernant l'ensemble des élèves français, les élèves scolarisés dans des établissements de l'éducation prioritaire et ceux de l'académie de Créteil sont issues de *Éducation prioritaire, Les réseaux "ambition réussite" et les réseaux de réussite scolaire, Académie de Créteil Données 2009-2010*, Repères pour l'analyse de gestion, Direction générale de l'enseignement scolaire, octobre 2010. Cette source ne couvre que le champ de l'enseignement public. La référence académie de Créteil a été reconstituée à partir des données désagrégées éducation prioritaire/hors éducation prioritaire de l'académie, en faisant l'hypothèse que la répartition éducation prioritaire/hors éducation prioritaire des élèves de troisième est identique à celle de l'ensemble des collégiens (pour lesquels la répartition est disponible).

b. Source : données DEPP. Les données disponibles pour l'académie de Créteil ne permettent pas de distinguer l'enseignement public de l'enseignement privé.

c. Sources : bulletins scolaires collectés dans les dossiers de candidature pour les moyennes. Données DNB du rectorat de Créteil pour les rangs. Les rangs sont calculés à partir des notes de français et de mathématiques obtenues au contrôle continu en situant l'élève candidat dans la distribution des notes de sa classe. Comme cette information est extraite des bases du brevet, elle n'est disponible que pour les élèves de troisième, dont les responsables légaux ont autorisé la divulgation des notes, et ne permet de situer les élèves candidats que parmi leurs camarades de classe dont les notes sont disponibles.

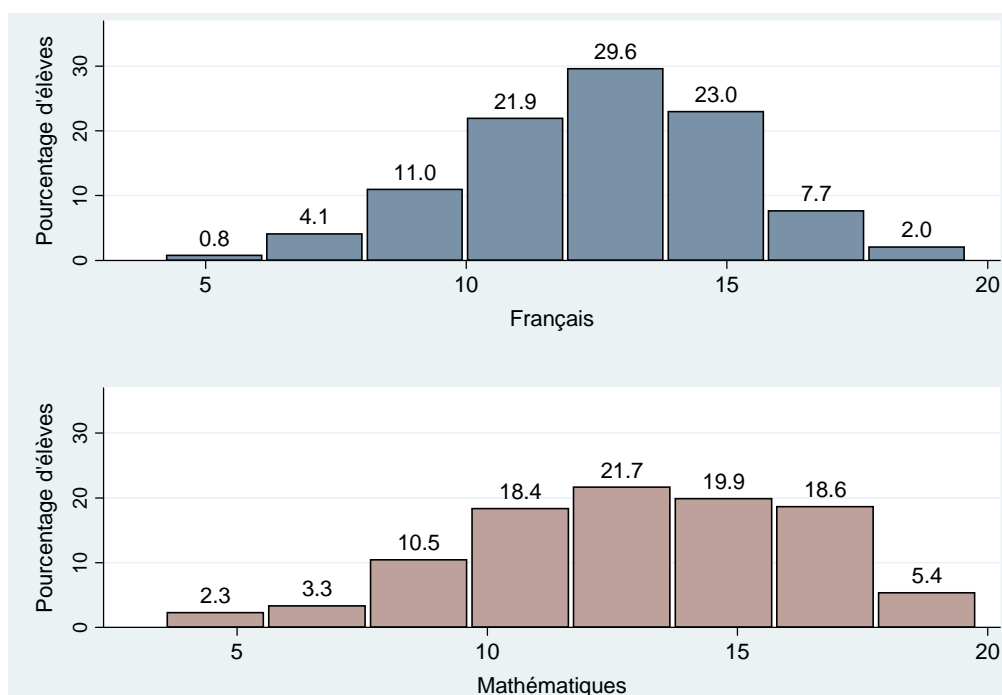


FIGURE 3 – Distribution des notes de contrôle continu avant la candidature

de classe. Si l'on fait l'approximation que ces notes suivent une loi normale¹⁶, on obtient que la moyenne au DNB des élèves retenus les place au 45ème percentile de la distribution de l'ensemble des élèves du pays. En d'autres termes, si on les comparait à 100 élèves tirés au hasard, on trouverait que la moyenne de leurs notes correspond à celle de l'élève classé 45ème sur cent, là où les notes de leurs camarades de classe d'origine correspondent à celle de l'élève classé 60ème.

Au total, le profil social des candidats retenus est bien celui d'élèves d'origine modeste, proches de la population scolarisée dans les établissements de l'éducation prioritaire. Ils s'en distinguent sans doute par des difficultés familiales et sociales particulières, qui sont un facteur important dans le processus de sélection des dossiers. Ces difficultés sont cependant des données confidentielles, que nous ne pouvons donc pas observer. Les élèves retenus, sans être les têtes de classe de leurs établissements d'origine, sont sensiblement meilleurs que leurs camarades, mais restent, au niveau national, des élèves proches de la médiane. Il est important de garder à l'esprit que l'intervention dont nous évaluons l'impact s'adresse à ce type d'élèves qui, loin d'être les élèves excellents que certains internats attendaient peut-être, ne sont pas non plus des élèves faibles. Rien ne peut être dit ici de ce que serait l'impact des internats

16. Il s'agit d'une approximation raisonnable pour ce type de mesure, en tout cas en dehors des queues de distribution.

d'excellence sur une population davantage en difficulté scolaire.

TABLE 5 – Autres statistiques descriptives sur les candidats

	Moyenne	N
Elèves candidats		
Fille	0.585	395
Age moyen	14.621	395
Famille des élèves candidats		
Famille monoparentale	0.369	352
Nombre d'enfants du responsable légal	2.880	393
Responsable légal sans diplôme	0.102	343
Responsable légal a le bac ou plus	0.283	343
Seul le français est parlé en famille	0.452	352
Décision de candidater à l'internat		
Idée du collège/d'un enseignant	0.482	353
Idée de l'enfant	0.244	353
Idée des parents	0.221	353

4 La scolarité à l'internat d'excellence de Sourdun

Les textes définissant la politique d'internat d'excellence laissent une marge de manœuvre importante aux équipes éducatives pour définir leur projet d'établissement et l'organisation de la scolarité. Ces équipes ont donc investi, chacune à leur manière, cette liberté qui leur était donnée (Rayou et Glasman, 2012). Dans ce qui suit, nous documentons les spécificités de la scolarité à Sourdun. Pour ce faire, nous comparons, sur un certain nombre de dimensions, le quotidien des élèves orientés vers l'internat à celui des élèves témoins, afin de déterminer dans quelle mesure ce dont les élèves bénéficient à Sourdun diffère de ce qu'on leur aurait offert dans l'établissement dans lequel ils seraient allés s'ils n'avaient pas été admis à Sourdun. Cela nous permet de répondre à différentes questions. Les cours à Sourdun se déroulent-ils de la même façon que dans les établissements témoins? Sourdun propose-t-il aux élèves un cadre de travail différent? Les professeurs, recrutés sur profil à Sourdun, diffèrent-ils de ceux qui enseignent dans les établissements témoins? Le cadre de l'internat change-t-il les relations entre les professeurs et leurs élèves? Ce cadre a-t-il un impact sur les relations entre élèves?

Une partie des données présentées ici sont issues des questionnaires soumis dans les mêmes conditions aux élèves de Sourdun et aux élèves témoins. Même si les questions étaient précises et souvent datées ("lundi dernier"), il ne s'agit pas de mesures objectives. Elles reflètent aussi la perception qu'ont les élèves de leurs conditions d'études et de vie.

4.1 L'emploi du temps et les conditions de travail à Sourdun

L'internat d'excellence met les équipes éducatives face à un défi nouveau : il ne s'agit plus de s'occuper des élèves huit ou neuf heures par jour, mais toute la journée. Ce défi est en même temps une opportunité : avoir toute la journée les élèves "sous la main" peut aussi être un moyen d'exercer sur eux une action plus forte. En outre, les ressources supplémentaires dont ont bénéficié l'internat d'excellence (via les moyens d'encadrement notamment), ainsi que le ciblage des élèves, sont susceptibles d'avoir modifié les conditions de travail en classe. Les tableaux qui suivent documentent ces changements.

Comme le montre la table 6, l'horaire de cours à Sourdun en français et en mathématiques suit les recommandations des programmes officiels. La plupart des différences qu'on peut observer proviennent du fait qu'à Sourdun les évaluations de contrôle continu n'ont pas lieu pendant les cours mais sur des plages horaires de deux heures dédiées. Ces heures sont ensuite déduites de l'horaire de cours de chaque matière, et les professeurs complètent leur service via de l'aide personnalisée, ou en encadrant des ateliers après les cours (atelier cuisine, activités artistiques, atelier astronomie, club journal, projet éco-internat...).

TABLE 6 – Heures hebdomadaires de français et de mathématiques

	Français		Mathématiques	
	Sourdun ^a	France ^b	Sourdun	France
Sixième	4h	4h30-5h	3h30	4h
Cinquième	3h30	4h	3h	3h30
Quatrième	3h30	4h	3h	3h30
Troisième	4h	4h30	3h30	4h
Seconde	3h30	4h	3h30	4h
Première S	4h30	4h	5h	5h
Première ES	5h	4h	3h	3h
Première L	5h	4h	2h	3h ^c

a. Source : Emplois du temps 2010-2011 à Sourdun.

b. Source : Site Internet du ministère de l'Éducation nationale ¹⁷

c. Uniquement si option.

Si l'horaire théorique est le même, la présence effective en cours est supérieure à Sourdun. Par exemple, seuls 35.8% des élèves déclarent avoir manqué les cours durant les 15 derniers jours, contre 51.6% dans le témoin. En outre, la proportion d'élèves qui déclarent avoir "séché" les cours au moins une fois est trois fois plus faible à Sourdun (cf. table 7, premier panel).

Le second panel de cette table montre que les conditions de travail en classe sont meilleures à Sourdun. Tout d'abord, les classes y sont plus petites (environ 19 élèves contre 24). Les cours

17. <http://www.education.gouv.fr/cid80/les-horaires-par-cycle-au-college.html> pour le collège; http://cache.media.education.gouv.fr/file/special_1/39/5/grilles_horaires_136395.pdf pour le lycée.

y semblent un peu moins troublés par le bruit et l'agitation, mais ces différences ne sont pas statistiquement significatives. Les élèves sont aussi moins nombreux à déclarer que les autres élèves n'écoutent pas ce que le professeur dit pendant les cours (16.1% contre 48.7% dans le groupe témoin).

TABLE 7 – Présence effective en cours et conditions de travail

	Moyenne Témoin	Sourdun - Témoin	σ	N
Présence effective en cours (15 derniers jours)				
Proportion d'élèves ayant manqué l'école	0.516	-0.158*	0.087	362
Proportion d'élèves ayant séché des cours	0.347	-0.233***	0.074	361
Conditions de travail en classe				
Nombre d'élèves de la classe	24.156	-4.757***	0.795	352
Le professeur doit attendre que les élèves se calment	0.389	-0.135	0.085	361
Les élèves ne peuvent pas bien travailler	0.245	-0.111	0.075	360
Il y a du bruit et de l'agitation	0.473	-0.135	0.093	361
Les élèves n'écoutent pas ce que dit le professeur	0.487	-0.326***	0.096	361

La table 8 compare l'emploi du temps en dehors des heures de cours des élèves de Sourdun à celui des élèves témoins. Le premier panel considère les devoirs, les études et les heures de soutien individuel avec un professeur. Les élèves de Sourdun consacrent plus de temps à faire leurs devoirs que les élèves témoins (6.8 heures par semaine contre 5 heures). Ils passent aussi trois fois plus de temps en études, et ils sont deux fois plus nombreux à bénéficier d'un soutien individuel de la part d'un de leurs professeurs.

Si on compare l'emploi du temps des deux groupes d'élèves le lundi (second panel), on constate que les élèves de Sourdun consacrent 0.8 heure de plus à leurs devoirs, et regardent 0.9 heure de moins la télévision. Le samedi (troisième panel), ils consacrent 0.6 heure de plus à leurs devoirs, et 0.5 heure de plus à faire du sport. Les élèves substituent donc des activités constructives (devoirs) à des activités non constructives (télévision) pendant la semaine. On n'observe pas de retour de balancier le week-end : les élèves scolarisés à Sourdun ne consacrent pas plus de temps à regarder la télévision ou à jouer aux jeux vidéos lorsqu'ils rentrent dans leurs familles le week-end. Au contraire, les habitudes prises la semaine semblent se prolonger le week-end, puisqu'ils consacrent plus de temps leurs devoirs le samedi aussi.

Enfin, les élèves bénéficient de plus d'activités culturelles (dernier panel). Au cours du dernier trimestre ils ont participé en moyenne à 1.5 sorties culturelles avec leur classe, contre 0.8 dans le témoin. Ils sont aussi plus de deux fois plus nombreux à être partis en voyage de classe. Ces voyages ne sont pas plus longs, mais ont plus souvent lieu à l'étranger.

TABLE 8 – Emploi du temps extra-scolaire

	Moyenne Témoin	Sourdun - Témoin	σ	N
Devoirs et études par semaine				
Heures consacrées aux devoirs	4.995	1.830**	0.765	355
Heures en étude ^a	1.909	3.478***	0.467	350
Soutien individuel ^b	0.090	0.105*	0.060	363
Heures en soutien individuel	0.810	0.584*	0.333	35
Emploi du temps le lundi				
Heures de devoirs	1.293	0.759***	0.179	356
Heures de lecture	0.399	0.255	0.179	353
Heures de sport	0.781	0.142	0.262	357
Heures de jeux vidéos	0.580	-0.122	0.219	352
Heures de télévision	1.415	-0.865***	0.262	356
Heure de coucher ^c	10.647	-0.277	0.305	363
Emploi du temps le samedi				
Heures de devoirs	1.602	0.584*	0.300	358
Heures de lecture	0.832	0.412	0.281	353
Heures de sport	1.410	0.534*	0.321	359
Heures de jeux vidéos	1.351	-0.029	0.412	349
Heures de télévision	2.633	0.255	0.472	354
Heure de coucher ^d	12.286	-0.321	0.343	363
Activités culturelles avec l'école				
Sorties (trimestre dernier)	0.842	0.635**	0.270	358
Voyage de classe cette année	0.273	0.397***	0.079	361
Durée du voyage (en jours)	5.955	-1.030	1.063	152
Voyage de classe à l'étranger	0.155	0.116*	0.063	361

a. La question était ainsi formulée : “La semaine dernière, combien de temps as-tu travaillé dans l'établissement en dehors des cours (étude surveillée, aide aux devoirs, CDI, soutien) ?”.

b. La question était ainsi formulée : “La semaine dernière, as-tu travaillé tout(e) seul(e) avec l'un de tes professeurs pour qu'il t'aide à faire tes devoirs ou à comprendre une leçon ?”.

c. On a enlevé (ajouté) 12 pour chaque heure déclarée avant (après) minuit.

d. Idem.

L'internat instaure donc un cadre qui limite les possibilités de “sécher” les cours ou de regarder la télévision, et réoriente une partie du temps des élèves vers leur travail scolaire. Il réduit également les problèmes de discipline en classe, même s'il est difficile de savoir ce qui tient à la mise en œuvre stricte d'un règlement et ce qui s'explique en amont par la sélection d'élèves moins perturbateurs. Outre cette dimension scolaire, l'internat propose aussi aux élèves une ouverture culturelle, via des sorties, des ateliers et des voyages. Ce tableau est très cohérent avec celui qui a été dressé par Rayou et Glasman (2012). Ces auteurs insistent notamment sur le règlement intérieur strict qui prévaut dans la plupart des internats qu'ils ont étudiés, et qui, par exemple, proscrit ou limite l'usage de la télévision. En outre, plusieurs établissements inclus dans leur étude ont aussi un projet d'ouverture culturelle, qui est une dimension importante de l'offre éducative des internats d'excellence.

4.2 L'équipe éducative de Sourdu et ses relations avec les élèves

L'équipe éducative de Sourdu a été recrutée sur profil. Il s'agit donc de professionnels volontaires et motivés par le projet, susceptibles de ce fait d'être très différents de leur collègues des établissements témoins. En outre, le cadre de l'internat modifie peut-être profondément les relations des professeurs avec les élèves. Nous documentons maintenant les spécificités de l'équipe éducative de l'internat, et les relations que les élèves développent avec elle¹⁸.

Le premier élément qui ressort est que le taux d'encadrement est supérieur à Sourdu. Le nombre d'heures de service de professeur par élève y est en effet 54% plus élevé que dans les établissements témoins (cf. table 9). Le ratio professeurs pour 100 élèves y est aussi 36% plus élevé. Ces différences proviennent principalement du fait que les classes sont plus petites à l'internat. Les écarts sont encore plus marqués lorsque l'on considère l'encadrement non pédagogique. Par exemple, le ratio de conseiller principal d'éducation (CPE) par élève est 3.7 fois plus élevé à Sourdu. Bien entendu, ce fort surcroît d'encadrement non pédagogique vient du fait que les élèves doivent être pris en charge jour et nuit pendant la semaine.

TABLE 9 – Encadrement des élèves

	Moyenne Témoin	Sourdu - Témoin	σ	N
Heures enseignant par élève	1.285	0.696***	0.033	373
Enseignants pour 100 élèves	8.198	2.974***	0.237	373
CPE pour 100 élèves	0.326	1.216***	0.018	375
Personnel administratif pour 100 élèves	1.444	4.169***	0.067	375
Assistants d'éducation pour 100 élèves	1.623	4.871***	0.113	375

18. La description des enseignants est faite sur la base de données administratives, tandis que celle des relations enseignants-élèves est faite à partir des questionnaires élèves.

Les professeurs à Sourdun sont plus diplômés mais moins expérimentés que leurs collègues des établissements témoins (cf. table 10). Ils sont en effet plus nombreux à être agrégés. Mais 14.1% d’entre eux ont moins d’un an d’expérience dans l’Éducation nationale contre 5.8% dans le témoin¹⁹. Aucun n’est en multi-affectation contre 17.5% dans les établissements témoins. Le recrutement sur profil semble donc avoir attiré de jeunes enseignants, un peu plus diplômés que la moyenne de ceux qui enseignent dans les établissements témoins.

TABLE 10 – Caractéristiques des enseignants par établissement

	Moyenne Témoin	Sourdun - Témoin	σ	N
Professeurs titulaires	0.802	-0.052	0.043	378
Professeurs agrégés	0.177	0.101***	0.020	378
Professeurs avec 0 à 1 an d’ancienneté	0.058	0.083***	0.007	378
Professeurs avec 1 à 3 ans d’ancienneté	0.136	0.119***	0.011	378
Ancienneté des professeurs	9.201	-3.493***	0.424	378
Professeurs en multi affectation	0.175	-0.180***	0.023	378
Professeurs à temps partiel	0.068	-0.047***	0.007	378

Les conditions d’enseignement dont ils bénéficient permettent aux enseignants de développer une pédagogie spécifique. S’ils semblent être un peu plus exigeants avec leurs élèves (table 11, premier panel), ce qui ressort de façon plus frappante est qu’ils semblent être mieux en mesure de se soucier de chacun d’entre eux (table 11, second panel). Les élèves de Sourdun sont par exemple beaucoup plus nombreux à déclarer que les professeurs s’intéressent aux progrès de chacun, qu’ils leurs permettent d’exprimer leurs opinions, ou qu’ils continuent à expliquer jusqu’à ce que tout le monde ait compris. On peut voir là les bénéfices de classes plus petites, plus homogènes en termes de niveau des élèves, et comptant moins d’éléments perturbateurs.

Le cadre de l’internat semble propice à l’instauration de relations de confiance entre élèves et professeurs (table 11, dernier panel). Par exemple, 90.5% des élèves de Sourdun déclarent bien s’entendre avec leurs professeurs, contre 63.1% dans le groupe témoin. Ils sont aussi plus nombreux à penser que s’ils en ont besoin leurs professeurs leur apporteront une aide supplémentaire. Ce climat favorable s’explique probablement par l’effet combiné des classes plus petites, des voyages de classe et du soutien individuel.

Les élèves ont en revanche de moins bonnes relations avec leurs CPE et leurs assistants d’éducation (table 12, premier panel). Comme l’ont noté Rayou et Glasman (2012), le rôle des CPE et des assistants d’éducation est très différent dans un internat d’excellence de ce qu’il est dans un établissement traditionnel. Ce sont eux qui deviennent “les référents de

19. La différence se reflète aussi dans le nombre moyen d’années d’expérience dans l’Éducation nationale (5.7 contre 9.2).

TABLE 11 – Pédagogie des professeurs, et relations avec les élèves

	Moyenne Témoin	Sourdun - Témoin	σ	N
Exigences des professeurs				
Le professeur contrôle les devoirs des élèves	0.437	0.085	0.083	360
Il est mécontent lorsque les élèves rendent un travail mal fait	0.615	0.146*	0.082	361
Il dit aux élèves qu'ils pourraient mieux faire	0.789	-0.031	0.067	361
Individualisation des rapports professeurs-élèves				
Le professeur s'intéresse aux progrès de chaque élève	0.486	0.270***	0.091	361
Il continue à expliquer jusqu'à ce que les élèves aient compris	0.466	0.409***	0.078	361
Il adapte la difficulté des devoirs au niveau des élèves	0.082	0.031	0.043	361
Il donne aux élèves l'occasion d'exprimer leurs opinions	0.492	0.305***	0.103	361
Les élèves n'osent pas demander quand ils n'ont pas compris	0.341	-0.037	0.078	361
Relations avec les professeurs				
Les élèves s'entendent bien avec les professeurs	0.631	0.274***	0.073	362
Les professeurs s'intéressent au bien-être des élèves	0.585	0.313***	0.084	362
Ils sont réellement à l'écoute	0.566	0.299***	0.093	362
Si besoin, ils apporteront une aide supplémentaire	0.676	0.256***	0.074	362
Ils traitent les élèves avec justice	0.562	0.224**	0.092	362

l'organisation du travail d'accompagnement scolaire". Ils sont aussi en charge d'organiser la vie de l'internat, et notamment d'y maintenir la discipline. C'est sans doute la raison pour laquelle les élèves ont de moins bonnes relations avec eux. Il est d'ailleurs intéressant de constater que les parents des élèves de Sourduin déclarent moins souvent être en conflit avec leurs enfants que les parents du groupe témoin (table 12, second panel). Le départ à l'internat pacifie les relations parents-enfants, comme l'ont aussi souligné Rayou et Glasman (2012) : "les relations avec les parents s'apaisent, ne serait-ce que parce que l'essentiel du travail scolaire se fait à l'internat, alors que c'est justement sur l'accomplissement du travail que les conflits éclatent facilement entre parents et enfants ou adolescents". Il semblerait donc que les conflits parent-enfant se transforment en conflits CPE-élève (ou assistant d'éducation-élève) lorsque l'enfant quitte le domicile parental pour l'internat.

4.3 Les relations entre élèves à Sourduin

Par ailleurs, l'internat est susceptible de modifier profondément la sociabilité des élèves. Le fait de vivre ensemble suscite-t-il le développement de relations d'amitié ou d'entraide entre élèves, ou exacerbe-t-il au contraire les tensions ?

Il ressort de nos données que l'internat favorise plutôt l'instauration d'un climat d'entraide entre élèves (table 13). Les élèves de Sourduin sont plus nombreux à déclarer que les bons élèves

TABLE 12 – Relations avec les CPE et les assistants d'éducation

	Moyenne Témoïn	Sourdun - Témoïn	σ	N
Relations avec les CPE/AE				
Les élèves s'entendent bien avec les CPE/AE	0.719	-0.294***	0.088	362
Les CPE/AE s'intéressent au bien-être des élèves	0.654	-0.015	0.099	362
Ils sont réellement à l'écoute	0.567	-0.066	0.103	361
Si besoin, ils apporteront une aide supplémentaire	0.558	-0.143	0.092	361
Ils traitent les élèves avec justice	0.654	-0.253***	0.095	362
Relations avec les parents				
Conflit enfant-parent	0.451	-0.199**	0.096	296

aident ceux qui ont de moins bonnes notes à faire leurs devoirs (59.8% contre 35.1%). Sur ce point, la vie à l'internat de Sourdun diffère de ce que Rayou et Glasman (2012) observent ailleurs : "Par ailleurs, les élèves ne semblent pas vraiment s'appuyer les uns sur les autres (...) en matière cognitive. À l'internat d'excellence, s'il arrive que l'on travaille avec les camarades, tout se passe comme si l'on ne pouvait apprendre qu'avec les enseignants ou les adultes qui encadrent". En outre, les élèves disent aussi avoir moins de difficultés à se faire des amis à l'école, même si cette seconde différence n'est pas statistiquement significative aux seuils usuels.

En outre, si l'on en croit les déclarations des élèves de notre échantillon, il est deux fois moins fréquent que des élèves commettent des actes de violence à l'internat que dans les établissements témoins. Cela renforce encore un peu l'image de l'internat qui se dégage des pages qui précèdent, comme un lieu dans lequel on met des élèves méritants "à l'abri" des éléments les plus perturbateurs de leurs établissements d'origine, pour leur permettre de se consacrer à leurs études.

TABLE 13 – Relations avec les autres élèves

	Moyenne Témoïn	Sourdun - Témoïn	σ	N
Relations avec les autres élèves				
Les élèves travaillent en groupe	0.340	0.093	0.086	361
Les bons élèves aident les moins bons	0.351	0.247***	0.086	360
Je me fais facilement des amis à l'école	0.754	0.123	0.076	362
Il y a souvent des conflits entre élèves	0.452	-0.034	0.090	362
Les élèves commettent des actes de violence	0.408	-0.186**	0.087	362

4.4 Sourdun : restaurer l'école ?

Si l'on voulait résumer les différences entre la scolarité dont les élèves bénéficient à Sourdun et celle qu'on leur aurait offerte ailleurs, on pourrait dire que Sourdun offre à ses élèves de bonnes conditions de travail, et un cadre qui leur permet de se concentrer sur leurs études. Sourdun ne propose pas à ses élèves une pédagogie fondamentalement différente, et l'horaire de cours y est sensiblement le même qu'ailleurs. Mais le régime de l'internat permet d'exercer un contrôle plus fort sur les élèves, qui ne peuvent plus (ou presque plus) sécher les cours, et qui consacrent désormais à leurs devoirs les heures qu'ils passaient jusque là devant la télévision. Les cours ne sont plus troublés par des problèmes de discipline voire de violence, permettant ainsi aux professeurs de se concentrer sur ce qui constitue le cœur de leur métier : faire progresser chaque élève. Ce constat est largement partagé par Rayou et Glasman (2012). Ils décrivent en effet la "ré-restauration d'un cadre" qui s'opère à l'internat, et citent à ce sujet les propos d'une mère d'interne : "Ce que je voulais d'un collège normal, on me le propose à l'internat d'excellence". Sourdun, ainsi que les internats d'excellence étudiés par Rayou et Glasman (2012), permettent donc simplement de restaurer les conditions d'une scolarité normale pour des élèves issus de quartiers populaires.

Il convient de préciser que tous les internats d'excellence ne fonctionnent pas sur ce modèle. Pour autant que nous ayons pu en juger lors de nos visites sur place, l'internat d'excellence de Montpellier offre à ses élèves un projet éducatif très différent, incluant notamment de nombreuses innovations pédagogiques qui reposent sur l'utilisation intensive des nouvelles technologies (horaire de cours adapté à chaque élève, suppression des notes, suivi hebdomadaire des progrès de chaque élève par un professeur référent à l'aide d'un tableau de bord des compétences etc.). Là où Montpellier essaie de changer les méthodes d'apprentissage des élèves, Sourdun cherche plutôt à restaurer un cadre propice au travail.

5 Le coût supplémentaire d'une année de scolarité à Sourdun

Le principal objectif de ce rapport est de mesurer l'impact de la scolarisation à Sourdun par rapport à une scolarisation dans un établissement traditionnel. Mais pour mettre en perspective cet impact, il est utile d'évaluer le coût additionnel qu'implique la scolarisation à Sourdun. Cela nous permettra en effet d'effectuer une analyse coût-efficacité : quelle est la plus-value de Sourdun, et quel est le coût de cette plus-value ?

La mesure du coût additionnel d'une année de scolarité à Sourdun par rapport aux alternatives que constituent les établissements où sont scolarisés les élèves témoins pose cependant deux difficultés. La première est technique : du fait du statut particulier de Sourdun et du fait que l'établissement était jusqu'à récemment en période de montée en charge, il n'existe pas

de données comptables strictement comparables entre Sourdun et les établissements témoins. Notre calcul doit donc reposer sur certaines hypothèses, que nous exposons de façon transparente dans ce qui suit. Nous proposons ainsi une fourchette permettant d'encadrer le coût supplémentaire d'une année à Sourdun, selon des hypothèses plus ou moins conservatrices²⁰.

La seconde difficulté porte sur les risques de mauvaise interprétation. Qui dit coût additionnel ne dit pas coût superflu. Il n'est en rien surprenant que la politique des internats d'excellence induise un coût additionnel, ne serait-ce que parce que la scolarisation en internat a pour contrepartie immédiate des coûts supplémentaires d'encadrement et d'hébergement. Quantifier ce coût additionnel ne sert pas à émettre un jugement normatif (est-ce trop cher ?), mais à permettre des comparaisons positives (pour un niveau d'impact donné, est-ce plus ou moins cher que des dispositifs alternatifs, comme par exemple la réduction de la taille des classes ?).

On pourrait être tenté d'aller au-delà de l'évaluation *globale* du coût additionnel de Sourdun. Quelle est dans ce coût la part qui s'explique par l'internat, et quelle est celle qui s'explique par un surcroît de dépenses pédagogiques (taux d'encadrement supérieur, activités culturelles) ? Nous devons cependant faire ici une claire mise en garde. D'une part, un tel exercice comporte nécessairement une part d'arbitraire : par exemple, les études surveillées dont les élèves bénéficient à Sourdun correspondent-elles à de la dépense pédagogique ou à des dépenses d'internat ? D'autre part, et de façon plus fondamentale, il faudra revenir au coût additionnel pris dans son ensemble pour "expliquer" la plus-value de Sourdun. En effet, rien ne permettra par exemple de "mettre de côté" le coût lié à l'internat et d'attribuer les meilleurs résultats des élèves au taux d'encadrement supérieur ou aux sorties culturelles : l'internat fait partie de l'intervention, et il est susceptible d'expliquer les effets observés tout comme les autres composantes du coût additionnel de Sourdun. Décrire ces composantes et les quantifier permet certes de se faire une idée de ce qu'est le dispositif ; mais cela ne permet pas d'attribuer les impacts observés à une composante plutôt qu'à une autre. Sourdun est un tout, ce tout a un coût et un impact que l'on peut évaluer ; aller au-delà serait spéculer.

Dans notre calcul du coût additionnel d'une année de scolarité à Sourdun, nous avons choisi de ne pas inclure les investissements initiaux nécessaires au lancement de Sourdun, et de nous en tenir autant que possible aux dépenses de fonctionnement *en régime de croisière*. Les raisons qui motivent ce choix sont à la fois d'ordre conceptuel et pratique. Tout d'abord, si la question à laquelle on souhaitait répondre était de savoir s'il faut fermer l'internat d'excellence de Sourdun, on ne devrait pas prendre en compte ces dépenses d'investissement dans notre calcul coût-efficacité, puisqu'elles constituent des coûts "irrécupérables" (*sunk costs*).

20. Le rapport en cours de la Cour des comptes permettra certainement des progrès appréciables par rapport aux calculs que nous présentons ici.

Les choses seraient différentes si on voulait évaluer l'opportunité d'ouvrir d'autres établissements sur le modèle de celui de Sourdun. Il alors faudrait prendre en compte les dépenses d'investissement, mais seulement à hauteur du surcroît nécessaire à la création d'un internat d'excellence par rapport à un établissement traditionnel. En effet, les investissements faits à Sourdun sont autant de dépenses en moins qu'il aurait fallu réaliser dans d'autres établissements scolaires. Le fait d'avoir construit Sourdun et d'autres internats d'excellence a créé de nouvelles places disponibles en collège-lycée, ce qui permettra un jour de ne pas ouvrir ou de ne pas rénover d'autres établissements. Calculer le surcroît d'investissement nécessaire à la création d'un internat d'excellence aurait été très compliqué. Il aurait en effet fallu être capable de n'inclure, dans l'ensemble des dépenses faites à Sourdun, que celles qui correspondent à la construction des bâtiments d'internat, car ce sont elles que l'on ne retrouverait pas dans un établissement traditionnel. Au total, un ratio coût-efficacité n'incluant que les dépenses de fonctionnement a un sens conceptuel clair, tandis qu'il nous aurait été difficile de construire un ratio clairement défini incluant aussi des dépenses d'investissement. Nous avons donc préféré nous en tenir au premier type d'indicateur.

Ces points clarifiés, la démarche suivie pour évaluer le coût additionnel induit par une année à Sourdun relève de la même logique que le reste de l'évaluation. On vise à comparer le coût de scolarisation des élèves de Sourdun avec ce qu'aurait été ce coût s'ils n'avaient pas été à Sourdun. Grâce au dispositif expérimental, cette comparaison serait immédiate si on disposait d'une mesure homogène des coûts de scolarisation à Sourdun et dans les établissements des élèves témoins. En pratique, un travail de reconstruction est nécessaire, car il n'existe pas de source permettant de calculer un coût moyen par élève au niveau de l'établissement²¹. On choisit de partir pour les établissements témoins du Compte de l'éducation publié par la DEPP²². Cette source, qui adopte l'approche exhaustive de la Comptabilité nationale, recense l'ensemble des dépenses d'éducation, quelle que soit leur modalité de financement (dotation de l'établissement, prestations fournies par les collectivités locales, personnel payé

21. Les comptes financiers de Sourdun et des établissements témoins ne suffisent pas pour cela : d'une part, ils ne couvrent pas les dépenses en personnel qui constituent la charge principale ; d'autre part, le périmètre des charges hors personnel du compte de Sourdun n'est pas le même que celui des autres établissements. Les établissements témoins sont en effet des établissements publics locaux d'enseignement (EPL) qui bénéficient d'un ensemble de prestations en nature de la part des collectivités locales (mise à disposition de personnel de cantine, transports scolaires, investissements dans les laboratoires de sciences, matériel informatique) qui ne figurent par conséquent pas dans leurs comptes. Au contraire, ces charges apparaissent dans les comptes de l'IE : Sourdun est un établissement d'État, qui bénéficie d'une subvention spécifique pour faire face à ces différentes charges, mais qui doit ensuite s'en acquitter lui-même. Celles-ci ne sont pas négligeables et ne peuvent être facilement isolées, si bien que la comparaison directe des comptes financiers n'a pas grand sens. Par exemple, pour rendre les dépenses de restauration de Sourdun comparables à celles d'un EPL, il faudrait enlever des dépenses de restauration à Sourdun les charges salariales du personnel de cantine, puisqu'elles sont prises en charge par la région dans un EPL et n'apparaissent donc pas dans les comptes. Ce n'est pas possible, dans la mesure où l'ensemble de la restauration à Sourdun est sous-traitée à un prestataire.

22. http://media.education.gouv.fr/file/199/02/3/DEPP-Dossier-2011-199-Compte-education_186023.pdf.

directement par l'État). Son principal inconvénient est de ne pas être disponible à un niveau plus désagrégé que celui de la région. On procède donc à une série d'ajustements pour tenir compte autant que possible des caractéristiques spécifiques des établissements témoins. Le coût moyen par élève témoin est ensuite mis en regard du coût total par élève à Sourdun, tel qu'il ressort du compte financier de l'internat d'excellence augmenté des charges salariales (qui ne transitent pas par le budget de Sourdun).

Le tableau 14 détaille les ajustements opérés pour rendre les deux coûts par élève obtenus aussi comparables que possible. On part des coûts moyens nationaux par élève, disponibles séparément pour le collège et le lycée (colonne a), dont on retranche les "coûts d'organisation du système"²³ ainsi que certaines activités mineures non couvertes dans les comptes de Sourdun, comme les leçons particulières. Le compte de l'éducation étant désagrégé par région²⁴, il est possible de corriger une éventuelle disparité entre le coût par élève en Île-de-France (où sont concentrés les établissements témoins) et le reste de la France. En pratique, la correction est quasi-nulle (coefficient de 1.04 et 1.00 au collège et au lycée, respectivement). Il se peut cependant qu'au sein de l'Île-de-France, les établissements dans lesquels se retrouvent les élèves témoins soient spécifiques. Hors charges de personnel, les comptes financiers des établissements fournis par la DEPP pour l'année 2010 permettent de noter des charges plus élevées dans les lycées témoins que pour la moyenne de l'académie de Créteil (coefficient de correction de 1.22 au lycée, pas de correction au collège). On choisit d'appliquer aussi cette correction aux dépenses en personnel, même s'il n'est pas possible d'estimer directement quelle serait la bonne correction. On obtient ainsi une mesure approchée du coût de scolarisation dans les lycées et collèges témoins (colonnes c et d). Pour obtenir le coût moyen qui aurait été celui des élèves scolarisés à Sourdun s'ils ne l'avaient pas été, on calcule enfin une moyenne pondérée qui tient compte de la part des collégiens à Sourdun (44% en 2010-2011). On parvient ainsi à un coût moyen reconstitué de 10 687 euros par élève.

Ce coût moyen est à comparer à celui de la scolarisation à Sourdun. On somme pour cela les dépenses du compte financier 2011 de Sourdun (rapporté au nombre moyen d'élèves sur la période) avec l'ensemble des dépenses salariales en 2011-2012 (rapportées au nombre d'élèves sur cette année scolaire). Le choix de l'année 2011, plutôt que des années 2009 ou 2010, sur lesquelles portent aussi l'évaluation, vise à s'approcher autant que possible d'un régime de

23. "Ces activités comprennent le pilotage et l'administration générale du système éducatif, l'exercice de la fonction de réglementation et de contrôle, la gestion centralisée extérieure aux établissements, par les services administratifs centraux, régionaux ou départementaux du ministère de l'éducation nationale, de la jeunesse et de la vie associative, mais aussi par les services administratifs au sein des collectivités territoriales par exemple. Les dépenses pour ces activités sont comptabilisées dans le domaine de l'éducation (c'est-à-dire, les dépenses aux niveaux académique et national)." (*Le Compte de l'éducation. Principes, méthodes et résultats pour les années 1999 à 2009*).

24. http://cache.media.education.gouv.fr/file/geo10/59/4/DEPP-geographie-ecole-2011_180594.pdf.

croisière, les premières années ayant donné lieu à d'importants investissements. Nous obtenons alors une dépense de 21 646 euros par élève. L'année 2011 reste malgré tout marquée par des investissements lourds correspondant encore au lancement de la structure. Ce premier chiffre surestime donc sans doute un peu la dépense par élève que l'on observera à Sourdun en régime de croisière. Par conséquent, nous calculons également un coût moyen par élève hors dépenses en capital (qui comportent essentiellement des investissements), et hors dépenses d'entretien (dont on peut penser qu'elles comportent une part de remise en état du site). On parvient ainsi à un coût de 19 045 euros par élève, qui sous-estime lui peut-être un peu la dépense en régime de croisière, puisque l'on observera alors des dépenses d'investissement et d'entretien même si elles seront moindres qu'en 2011.

On obtient ainsi une estimation du coût additionnel de la scolarité à Sourdun qui se situe entre 8 000 et 11 000 euros. La dépense par élève à Sourdun est donc environ deux fois supérieure à celle observée dans les établissements témoins. Le coût en personnel est environ 50% plus élevé, ce qui traduit une plus forte présence des enseignants et plus encore un encadrement renforcé en vie scolaire, qui ont été documentés dans les tableaux 9 et 10. Les autres dépenses sont multipliées par 2.5, ce qui reflète le fonctionnement de l'internat, mais aussi d'autres spécificités du projet pédagogique de Sourdun, sans que nous soyons en mesure de séparer précisément ces deux éléments. C'est ce surcroît de dépenses qui, pris dans son ensemble, doit être mis en regard des impacts de l'internat d'excellence, détaillés dans les sections suivantes.

6 L'impact de la scolarité à Sourdun sur les élèves bénéficiaires

6.1 Impact sur le niveau scolaire des élèves

Cette section présente l'effet d'une scolarité à Sourdun sur les résultats à des tests standardisés en français et en mathématiques²⁵. Les mêmes tests ont été soumis chaque fin d'année pendant deux ans à l'ensemble des élèves, scolarisés ou non à Sourdun, comme l'explique la section 2.4. Ils permettent donc, dans ces deux matières, de mesurer les différences de compétences scolaires que l'on on peut attribuer à l'internat. L'utilisation de tests standardisés est importante car les notes données par les enseignants dans les différentes matières n'apportent aucune information utile si les pratiques de notation diffèrent entre Sourdun et les autres établissements. Dans cette section, nous utilisons également les résultats au diplôme national du brevet (DNB), qui apportent un complément d'information, bien qu'ils soient partiellement

25. La focalisation sur le français et les mathématiques, bien que fréquente dans ce type d'études, n'est pas entièrement satisfaisante ; il était cependant difficile d'étendre les interrogations à d'autres disciplines, puisque cela aurait supposé de mobiliser les élèves sur de nombreuses plages horaires.

TABLE 14 – Coût par élève à Sourdun et dans les établissements témoins

	France			Etablissements témoins			Sourdun	
	1er cycle (a)	2nd cycle (b)		1er cycle (c)	2nd cycle (d)	Moyenne pondérée (e)	Total (f)	Total hors opérations en capital et entretien (g)
Coût moyen par élève (0)	7,613	10,281		7,908	12,571	10,687	21,646	19,045
dont :								
dépenses de personnel (1)	5,996	8,097		6,228	9,900	8,393	12,167	12,167
autres dépenses (2)	1,617	2,184		1,680	2,670	2,295	9,479	6,878

Cellules a0 et b0 : coût moyen par élève dans le secondaire hors coûts d'organisation du système, des leçons particulières, d'habillement et autres (source : Compte de l'éducation (2011), tableaux 1, 8 et 10).

Cellules c1-c2 et d1-d2 : coûts nationaux (a2 et b2) corrigés du facteur Île-de-France/moyenne nationale (source : Géographie de l'école) et du facteur établissements témoins/Île-de-France (source : comptes financiers des établissements, 2010).

Cellules e0-e2 : moyenne pondérée des colonnes c et d en fonction de la part du collège et du lycée à Sourdun.

Cellules f0-g2 : coûts moyens par élève, issus du compte financier de Sourdun (2011) et de la masse salariale 2011-2012.

Les données pour la France portent sur l'année 2008. Elles sont ajustées pour tenir compte de la hausse du point d'indice (dépenses en personnel) et de l'inflation (dépenses hors personnel) entre 2008 et 2011. Les données pour Sourdun portent sur l'année 2011.

déterminés par les notes de contrôle continu.

6.1.1 Les tests standardisés et les scores construits par la DEPP

Les tests standardisés ont été conçus par la Direction de l'évaluation, de la prospective et de la performance (DEPP) du ministère de l'Éducation nationale. Des batteries d'épreuves de français et de mathématiques couvrant tous les niveaux de la 6^{ème} à la 1^{ère} ont été produites spécifiquement pour cette évaluation. Elles s'appuient sur des tests existants mais les complètent pour pouvoir les déployer sur l'ensemble des niveaux du second degré. Les tests ont été (sauf exception) soumis en ligne (cf. section 2.4). Ils se présentaient sous forme de questionnaires à choix multiples. En français, il s'agissait pour l'essentiel de lire des textes (extraits d'articles de journaux, de romans et de pièces de théâtre) et de répondre à des questions s'y rapportant. Ces tests visaient à mesurer la capacité des élèves à comprendre un texte et à en extraire de l'information. Ils s'inscrivent donc dans une tradition de mesure de "*littératie*", dont relèvent notamment les questionnaires PISA. Il n'y avait pas de production écrite. Les tests de mathématiques comportaient une trentaine d'exercices mettant en jeu des notions d'algèbre et de géométrie au programme du niveau concerné. À aucun moment, les enseignants n'ont pu disposer des tests, dont le contenu a toujours été tenu secret.

À l'issue de la première session de tests, la DEPP a apporté des corrections et des ajustements à leur contenu. Après la dernière session de tests, qui a eu lieu en juin 2012, les données ont été transmises par l'École d'économie de Paris à la DEPP, afin que celle-ci construise pour chaque élève une mesure synthétique de sa réussite aux tests, séparément en français et en mathématiques. On évalue ensuite l'impact de la scolarisation à Sourdun sur cette mesure.

La DEPP a effectué ce travail à l'aveugle, c'est-à-dire sans savoir quel élève appartenait au groupe orienté vers l'internat et quel élève appartenait au groupe témoin (ni quel élève était effectivement à l'internat), ce qui garantit que les scores ont été construits de façon objective, sans privilégier, de façon consciente ou inconsciente, un groupe par rapport à l'autre. Pour construire les scores de mathématiques et de français, la DEPP a utilisé un modèle dit de "réponse à l'item", fréquemment utilisé en psychométrie. Ce modèle est présenté dans l'Annexe de ce rapport²⁶.

Il existe de multiples façons de mesurer les compétences d'élèves en français et en mathématiques. Chaque jeu de tests ou chaque examen ne permet de mesurer qu'un certain nombre de savoirs et savoir-faire, et omet donc certaines dimensions qui pourraient être pertinentes.

26. Les résultats présentés ici sont peu différents si on utilise le score brut des élèves à chacun des tests (leur pourcentage de bonnes réponses) au lieu des mesures construites par la DEPP. Néanmoins, ce traitement a permis de retirer des scores certains items qui sont apparus peu fiables. Ce travail a donc permis de construire des mesures de compétences purgées d'une certaine forme d'aléa statistique (élèves qui répondent au hasard à certaines questions qu'ils ne comprennent pas, etc.).

C'est indéniablement le cas de nos tests, d'autant plus que la passation en ligne a imposé certaines contraintes sur le type de questions posées. Aussi, afin de vérifier que nos tests sont corrélés à d'autres mesures possibles, nous avons examiné leur corrélation avec les notes des épreuves terminales du brevet, en utilisant l'échantillon des élèves pour lesquels les deux mesures sont disponibles simultanément.

La table 15 présente ainsi les coefficients de corrélation entre les scores de français et de mathématiques construits par la DEPP et les notes de français et de mathématiques (respectivement) à l'épreuve terminale du brevet. Par construction, un coefficient de corrélation entre deux variables est compris entre -1 et 1 , 1 signifiant que la corrélation entre les deux variables est parfaite et positive, tandis que -1 signifie que cette corrélation est parfaite et négative. Ici, les coefficients que l'on obtient, 0.48 et 0.59 , sont positifs et peuvent être considérés comme élevés dans la mesure où les deux notes ne sont pas basées sur le même type d'exercice. Pour avoir un point de comparaison, on peut considérer la corrélation entre les notes de contrôle continu du brevet et les notes à l'épreuve terminale du brevet. On obtient alors des corrélations légèrement plus élevées entre ces deux types de notes (0.6 en français et 0.62 en mathématiques), mais d'ordre de grandeur comparable, voire extrêmement proche en mathématiques. Cela signifie donc que les scores construits par la DEPP sont très bien corrélés aux notes des élèves au brevet et qu'ils ne mesurent pas des compétences qui seraient très spécifiques et sans rapport avec des notes habituellement utilisées dans le système scolaire.

TABLE 15 – Corrélation entre les notes au brevet et les scores au test

	Corrélation DNB/Scores	N
Français	0.48	217
Mathématiques	0.59	215

6.1.2 Au bout de deux ans, Sourdu produit un fort effet en mathématiques

La table 16 présente l'effet moyen, à travers tous les niveaux, de chaque année de scolarisation à Sourdu sur les résultats des élèves aux tests de français et de mathématiques. Le score construit par la DEPP est normalisé pour avoir une moyenne de 0 et un écart-type de 1 , si bien que ces effets s'interprètent directement en pourcentage d'écart-type. Ces effets sont estimés à partir de 744 observations pour les tests de français et 735 pour les tests de mathématiques, car pour chaque élève on utilise les deux observations correspondant à ses résultats aux tests passés un an puis deux ans après sa candidature. Les coefficients présentés dans cette table s'interprètent comme l'effet causal de la première année passée à Sourdu,

puis l'effet *additionnel* d'une seconde année²⁷. L'effet total au bout de deux ans est la somme de ces deux coefficients.

TABLE 16 – Effet par année à Sourdun sur les scores cognitifs

Matière	Effet d'une 1 ^{ère} année à Sourdun	σ	Effet d'une 2 ^{ème} année à Sourdun	σ	N
Français	-0.051	0.147	-0.094	0.173	744
Mathématiques	-0.003	0.128	0.411**	0.173	735

De cette table, on peut tirer deux conclusions. D'abord, une première année de scolarisation à Sourdun n'a pas d'effet sur les scores des élèves, ni en français, ni en mathématiques. Les coefficients sont très petits (moins de 5% d'un écart-type de score) et surtout ils ne sont pas statistiquement significatifs et sont donc uniquement dus aux aléas d'échantillonnage. De ce point de vue, leur signe négatif ne signifie pas grand chose : il n'y a simplement pas de différence visible entre les élèves de Sourdun et les autres. Deuxième conclusion, une seconde année à Sourdun n'a pas non plus d'effet sur le score des élèves en français, mais a un fort effet sur leurs résultats en mathématiques, de l'ordre de 40% d'un écart-type du score. Cette quantification peut sembler abstraite, mais elle est couramment utilisée dans les recherches qui évaluent des progressions scolaires afin de rendre comparables des résultats issus de diverses études. En pratique, l'effet que nous obtenons ici signifie par exemple qu'un élève initialement médian (c'est-à-dire qui était classé 50ème sur 100) atteint en moyenne, grâce à l'intervention, le niveau de celui qui était classé 35ème sur 100. Dans les études qui portent sur les effets des politiques scolaires, un impact de cet ordre est considéré comme élevé ; peu d'interventions permettent d'obtenir de tels progrès (nous apportons des points de comparaison plus loin).

C'est donc au bout de deux ans seulement que Sourdun produit des effets sur les compétences des élèves. Il semble que ces derniers passent d'abord par une phase d'adaptation voire de découragement lorsqu'ils arrivent à l'internat. Nous avons mesuré la perception que les élèves ont de leurs compétences scolaires au moyen de la version française du *Self Perception Profile for Adolescents* (Bouffard, Seidah, Mcintyre, Boivin, Vezeau, et Cantin, 2002). Cette échelle psychométrique mesure l'estime de soi scolaire des élèves à partir d'un grand nombre de questions, par exemple en leur demandant de dire quelle affirmation leur correspond le mieux entre : "Certains jeunes sont très lents à finir leur travail scolaire", et "D'autres jeunes peuvent faire leur travail scolaire rapidement". Le tableau 17 présente l'effet de la scolarité à

27. Note technique : ces coefficients sont estimés par la méthode des variables instrumentales, qui généralise la méthode présentée dans la section 2.2 au cas où il y a deux paramètres à estimer simultanément. Les variables d'intérêt valent respectivement 1 si l'élève a passé une année à Sourdun, et 0 sinon ; et 1 si l'élève a passé deux années à Sourdun, et 0 sinon. Ces deux variables sont instrumentées par une indicatrice pour l'orientation vers Sourdun, interagie avec une indicatrice pour la passation du test la première année et avec une indicatrice pour la passation du test la deuxième année.

l'internat sur cette mesure. L'effet après un an à Sourduin est négatif, tandis qu'une seconde année a un effet positif, même s'il convient de noter que seul ce deuxième effet est significatif. Chez les filles, les lycéens, et les élèves initialement les plus faibles²⁸, la baisse de l'estime de soi scolaire lors de la première année d'internat est statistiquement significative. Rayou et Glasman (2012) dressent le même constat. Lors des entretiens réalisés par ces auteurs, plusieurs internes évoquent le "choc" de se retrouver confrontés, à l'arrivée en internat d'excellence, à des exigences et des méthodes d'enseignement nouvelles, et à un groupe de pairs de meilleur niveau scolaire. Les auteurs observent aussi que les élèves enregistrent une nette baisse de leurs notes par rapport à l'année précédente lors du premier trimestre. L'entrée en internat conduirait donc certains élèves à perdre confiance en eux dans un premier temps, et ce n'est qu'une fois cette phase de découragement surmontée que les effets positifs commencent à émerger.

TABLE 17 – Effets sur l'estime de soi scolaire

	Moyenne Témoin	Effet d'une 1 ^{ère} année à Sourduin	σ	Effet d'une 2 ^{ème} année à Sourduin	σ	N
Effet sur la population						
	0.027	-0.199	0.145	0.364*	0.192	735
Effets par sous-groupes						
Filles	0.028	-0.392**	0.157	0.332*	0.187	430
Garçons	0.056	0.162	0.302	0.354	0.265	305
Collégiens	-0.020	0.009	0.186	0.554**	0.235	353
Lycéens	0.088	-0.374*	0.209	0.203	0.210	382
Forts F	0.306	-0.055	0.205	0.205	0.195	380
Faibles F	-0.263	-0.343*	0.195	0.546**	0.245	350
Forts M	0.330	-0.135	0.217	0.142	0.205	383
Faibles M	-0.260	-0.255	0.195	0.598***	0.227	347

L'absence d'effet mesuré en français peut surprendre. Il convient d'abord de souligner que ce résultat peut être lié à la taille modeste de notre échantillon, qui ne permet pas de détecter de différences fines entre les groupes. L'expérience est calibrée pour révéler des impacts avec une forte probabilité s'ils atteignent au moins 35% d'un écart-type du score (ce qui est le cas en mathématiques). On a considéré que l'investissement dans l'internat d'excellence de Sourduin était suffisamment important pour que le protocole expérimental soit seulement capable de détecter des effets assez grands. S'il existe des effets en français, mais qu'ils sont faibles,

28. Pour construire ces sous-groupes, nous avons utilisé les bulletins scolaires fournis par les élèves dans leur dossier de candidature. Les élèves faibles sont définis comme ceux en dessous de la médiane de leur niveau scolaire dans la matière concernée.

par exemple 15% d'un écart-type, ils peuvent ne pas être apparents avec l'échantillon d'un peu moins de 400 élèves dont nous disposons²⁹. En d'autres termes, on ne peut pas faire la distinction entre l'absence d'effet et un effet positif mais relativement modeste.

Il faut aussi souligner qu'il est fréquent, lorsque des évaluations rigoureuses d'interventions scolaires sont menées, d'obtenir des impacts forts en mathématiques et sensiblement plus faibles ou non-significatifs en langue, même si ce n'est pas systématique. Nous mettons plus loin ces résultats en perspective avec les évaluations de programmes à l'étranger, dont certaines dimensions sont comparables à celles de l'internat d'excellence. Dans ces études, il est fréquent d'obtenir des impacts notables en mathématiques et des impacts nuls ou beaucoup plus faibles en anglais (Abdulkadiroglu, Angrist, Dynarski, Kane, et Pathak, 2011, Angrist, Dynarski, Kane, Pathak, et Walters, 2010, Dobbie et Fryer, 2011). Il semble donc qu'il est plus difficile d'accélérer les apprentissages en langue qu'en mathématiques, surtout chez des adolescents. En effet, certains travaux en psychologie du développement montrent que les aptitudes verbales des enfants sont fixées plus tôt dans la vie que leurs aptitudes numériques et mathématiques (cf. Hopkins et Bracht, 1975).

6.1.3 Ces effets sont cohérents avec les évaluations de dispositifs similaires

Il convient d'abord de comparer ces résultats à ceux d'études portant sur des dispositifs similaires. Cet exercice est difficile car les internats d'excellence constituent une politique originale. En France, on ne dispose pas d'études portant sur un programme comparable. Les *charters schools* américaines sont sans doute le programme qui présente le plus de similitudes avec Sourdun. Il s'agit en effet d'établissements d'enseignement secondaire accueillant des élèves de milieux défavorisés, qui doivent respecter des codes comportementaux stricts et signer un contrat par lequel ils s'engagent à tout mettre en oeuvre pour atteindre l'excellence (modèle du "*no excuses*" charter school). Les journées scolaires y sont plus longues, notamment en raison d'un encadrement des devoirs, et les professeurs souvent plus jeunes que dans une école publique traditionnelle. La dépense par élève est variable d'une charter school à l'autre, mais les ordres de grandeur sont comparables à la dépense par élève à Sourdun³⁰. Une différence notable est que les *charters schools* ne comportent en général pas d'internat, à l'exception de la "*SEED Boarding School*" près de Washington D.C..

Ces *charters schools* ont fait l'objet de plusieurs évaluations par expérimentation contrôlée.

29. Dans ce cas, l'effet, non-significatif, que nous obtenons dans la table peut prendre des valeurs assez indéterminées, généralement petites, et négatives ou positives, car elles sont guidées par l'aléa d'échantillonnage; c'est bien ce qui se passe ici.

30. Par exemple, la dépense par élève dans la Promise Academy School de Harlem est de 19 000 dollars (15 000 euros) par an. Dans la "*SEED Boarding School*" près de Washington D.C., cette dépense s'élève à 39 000 dollars (30 000 euros) par an en raison de l'internat.

Les effets observés à Sourdun sont extrêmement proches de ceux produits par la Promise Academy School à Harlem, un quartier défavorisé de New-York (Dobbie et Fryer, 2011). Chaque année de collège passée dans cette école accroît en effet les scores des élèves en mathématiques de 23% d'un écart-type de score, soit 46% au bout de deux ans ; en revanche, l'effet sur les scores en anglais est très légèrement positif, mais non significativement différent de 0. Les *charters schools* de Boston étudiées par Abdulkadiroglu, Angrist, Dynarski, Kane, et Pathak (2011) et Angrist, Dynarski, Kane, Pathak, et Walters (2010) produisent des effets sensiblement supérieurs à ceux observés à Sourdun, mais qui sont là encore beaucoup plus forts en mathématiques qu'en anglais (+35% d'écart-type au bout d'un an en mathématiques, et +12% en anglais). Enfin, l'évaluation de Curto et Fryer (2011) porte sur la "SEED Boarding School", un dispositif comparable à Sourdun en raison du régime d'internat, de sa cible de jeunes d'origine défavorisée et de ses fortes exigences à l'égard des élèves. Les auteurs trouvent un effet de 20% d'écart-type en anglais par année passée dans cette école au niveau collège et 23% en mathématiques. Cette mise en perspective montre que les résultats que nous obtenons ici sont assez cohérents avec ceux que l'on trouve dans les travaux de recherche qui portent sur les dispositifs les plus comparables.

On peut noter un dernier élément intéressant. Curto et Fryer (2011) concluent leur étude en affirmant que les *charters schools* sans internat ont un rapport coût-bénéfice meilleur que celui de la "SEED Boarding School". En effet, celle-ci ne produit pas des effets beaucoup plus forts que ceux des autres *charters schools*, bien que le régime de l'internat accroisse fortement son coût. Il n'est pas possible d'affirmer que ce résultat vaudrait aussi en France, mais cela conduit tout de même à s'interroger sur la contribution spécifique du régime d'internat par rapport aux autres dimensions de Sourdun (recrutement des élèves et des enseignants, encadrement des devoirs, aide individualisée et projet pédagogique).

6.1.4 L'internat d'excellence de Sourdun comparé à des politiques alternatives

On peut également mettre en perspective les effets produits par Sourdun avec ceux que l'on obtiendrait en déployant d'autres politiques. Le champ des politiques scolaires est extrêmement vaste. Certaines sont systémiques : les chercheurs discutent ainsi des effets des modes de gouvernance (mise en concurrence des écoles, contrôle de leurs performances, marché du travail des enseignants), mais il existe peu de consensus dans ces domaines. D'autres études considèrent l'effet des méthodes pédagogiques, des pratiques des enseignants et de la formation des maîtres. Elles ont permis d'identifier plusieurs outils pédagogiques qui se sont révélés efficaces lorsqu'ils ont été expérimentés à petite échelle. Néanmoins, il s'avère souvent délicat de les mettre en œuvre à grande échelle, car cela suppose de modifier en profondeur les pratiques des enseignants ; la littérature qui estime les effets de la formation institutionnelle des

professeurs sur les résultats des élèves est peu fournie et les effets estimés sont assez souvent décevants.

Une troisième catégorie d'actions consiste à augmenter les ressources mobilisées au profit des élèves. Les internats d'excellence appartiennent à cette catégorie : ils concentrent des ressources importantes sur un faible nombre de collégiens et de lycéens, pour restaurer de bonnes conditions de travail pour ces élèves. Mais il existe d'autres façons d'utiliser des ressources. On peut par exemple les répartir sur un nombre plus important d'élèves. La politique d'éducation prioritaire traditionnelle constitue un bon exemple de ce type de politiques : elle touche en effet un nombre très important d'élèves, mais le surcroît de ressources effectivement mobilisé par élève est très faible (voir Bénabou, Kramarz, et Prost, 2004, Beffy et Davezies, 2012 et Cour des comptes, 2010). La seule évaluation quantifiée disponible ne trouve aucun impact de cette politique sur la réussite des élèves (Bénabou, Kramarz, et Prost, 2004). En ce sens, l'évaluation de l'internat d'excellence de Sourdun montre qu'il est possible de faire progresser les élèves de groupes prioritaires, à condition de concentrer effectivement sur eux un surcroît conséquent de ressources.

Dans cette section, nous mettons en perspective l'internat avec deux autres interventions qui consistent elles aussi à augmenter de façon importante les ressources mobilisées au profit des élèves, et qui ont été évaluées dans la littérature. La première politique alternative que nous considérons consiste à cibler des enfants beaucoup plus jeunes, en intervenant dès le primaire ou même avant. La seconde consiste à diminuer la taille des classes. Cette intervention augmente la principale ressource du processus éducatif : le temps d'enseignant par élève, là où l'internat joue sur plusieurs de ces ressources (enseignants, CPE, assistants d'éducation, sorties culturelles, heures d'études...).

Interventions précoces ou remédiation ?

Des interventions intensives visant de très jeunes enfants d'origine défavorisée ont été expérimentées aux États-Unis. Elles mêlent stimulation des enfants et sensibilisation des parents aux enjeux éducatifs et de santé. Elles reposent sur l'idée que les processus cognitifs et non cognitifs sont particulièrement malléables aux jeunes âges, si bien que des interventions précoces pourraient être plus efficaces que des actions de remédiation plus tardives.

Plusieurs de ces interventions ont été évaluées à l'aide de protocoles randomisés. Certaines ont des effets spectaculaires, encore visibles à l'âge adulte. Par exemple, à 27 ans, 83% des femmes qui ont bénéficié du programme Perry Preschool lorsqu'elles avaient entre 3 et 5 ans ont un emploi, contre 55% dans le groupe témoin. Le coût de ces dispositifs (typiquement 10 000 dollars par an) est comparable à celui d'une année de scolarisation (Currie, 2001). Sur la base de ces études, certains chercheurs ont comparé le ratio coût-bénéfice de ces interventions

précoces à celui de politiques de remédiation, et concluent qu'aux États-Unis les interventions précoces sont plus efficaces que les interventions postérieures (voir Knudsen, Heckman, Cameron, et Shonkoff, 2006 et Cunha, Heckman, Lochner, et Masterov, 2006).

Cependant, les dispositifs de remédiation inclus dans ces comparaisons coût-bénéfice sont très différents de l'intervention proposée à l'internat d'excellence de Sourdun (voir Cunha, Heckman, Lochner, et Masterov, 2006, Tableau 9). Par exemple les *charter schools* ne sont pas incluses, faute de disposer d'un recul suffisant sur leurs effets à l'âge adulte. À ce stade, il n'est donc pas encore possible de dire si aux États-Unis, le ratio coût-bénéfice des *charter schools* est meilleur ou non que celui des interventions précoces. En outre, il nous est impossible de comparer les coûts et les bénéfices produits par l'internat d'excellence de Sourdun à ceux produits par une intervention précoce en France, où aucune intervention de ce type n'a fait l'objet d'une évaluation rigoureuse. À ce jour, nous disposons d'un seul élément qui tend à montrer qu'en France aussi les interventions sont plus efficaces lorsqu'elles sont précoces. Piketty et Valdenaire (2006) trouvent en effet que des réductions de taille de classe produisent des effets deux fois plus forts au primaire qu'au collège.

En résumé, loin de pouvoir affirmer que des interventions précoces seraient plus efficaces que l'internat d'excellence de Sourdun, nous devons nous contenter de rappeler qu'une littérature importante a montré la grande efficacité de certaines de ces interventions aux États-Unis, et que la cible d'âge sur laquelle porte un dispositif est un choix fondamental.

Construire des internats ou réduire la taille des classes ?

Une autre façon d'améliorer les conditions de travail des élèves et de favoriser l'instauration de relations de confiance avec leurs professeurs est de réduire la taille des classes. Il semble donc intéressant de comparer les effets de cette politique à ceux de Sourdun. Ce point de comparaison est d'autant plus intéressant que son coût est assez simple à calibrer. Puisque la dépense par élève induite par l'internat de Sourdun est environ deux fois supérieure à celle des établissements dans lesquels les élèves témoins sont scolarisés, on peut comparer les impacts que nous obtenons avec l'effet d'une réduction de moitié de la taille des classes, puisqu'elle conduirait aussi à doubler à peu près la dépense par élève³¹.

Pour ce faire, nous pouvons nous appuyer sur l'analyse de Piketty et Valdenaire (2006), qui ont étudié l'impact d'une réduction de la taille des classes sur les résultats scolaires des

31. Comme on l'a vu dans la section 5, la dépense par élève à Sourdun est entre 78% et 102% supérieure à celle que l'on observe dans les établissements témoins, selon que l'on choisit des hypothèses plus ou moins conservatrices. Diviser la taille des classes par deux conduirait mécaniquement à doubler le nombre d'enseignants, et à augmenter le nombre de CPE et d'assistants d'éducation (augmentation du nombre d'études à surveiller, des réunions parents-professeurs à organiser...). Au total, cette politique multiplierait presque par deux les dépenses de personnel, qui représentent environ 80% de la dépense par élève en France (cf. section 5). En outre, elle induirait sans doute d'autres dépenses de réorganisation, si bien qu'au total on peut considérer que les deux politiques ont un coût à peu près équivalent.

élève³². Cette comparaison est naturellement indicative. Tout d’abord cette évaluation ne porte pas sur les mêmes populations que celle que nous étudions ici. Ensuite, l’étude de Piketty et Valdenaire (2006) et la nôtre sont soumises à un aléa d’échantillonnage ; les chiffres reportés ne sont donc pas le “vrai” effet d’une réduction de la taille des classes ou d’une scolarité à l’internat d’excellence de Sourdu, mais une estimation qui s’accompagne d’une marge d’erreur statistique. Les écarts entre les effets de ces deux politiques ne doivent donc pas être surinterprétés. Il faut surtout retenir de cet exercice les ordres de grandeur et se demander si ces politiques ont des effets comparables, ou si leurs effets sont clairement différents.

Passer deux années à Sourdu produit un effet de 40.8% d’écart-type sur le score en mathématiques des élèves, soit 20.4% par année, tandis que l’effet sur le score en français n’est pas significativement différent de 0. On obtient donc un effet cumulé en français et en mathématiques de $20.4 + 0 = 20.4\%$ d’écart-type par année passée à Sourdu, si on donne le même poids aux deux matières. Cet effet est estimé sur une population de collégiens et de lycéens socialement défavorisés, mais dont le niveau scolaire initial était comparable à celui de la moyenne des élèves français (cf. Table 4).

Piketty et Valdenaire (2006) trouvent qu’une diminution d’un élève de la taille des classes pendant une année conduit en moyenne à une augmentation de 0.8% d’écart-type de la note des élèves à l’épreuve terminale de mathématiques du Brevet. L’effet sur le résultat à l’épreuve de mathématiques du Baccalauréat n’est pas significativement différent de 0. En français, réduire d’un élève l’effectif augmente de 1.1% d’écart-type la note à l’épreuve terminale du Brevet. L’effet sur la note au Baccalauréat en français est non significativement différent de 0 (cf. Piketty et Valdenaire, 2006, Tableau 18). Étant donné que notre échantillon contient à peu près le même nombre de collégiens que de lycéens, et en donnant à nouveau le même poids au français et aux mathématiques, on obtient un impact total de $\frac{0.8+1.1}{2} = 0.95\%$ d’écart-type par élève en moins. Les élèves du groupe témoin sont en moyenne scolarisés dans des classes de 24 élèves (cf. Table 7). Diviser la taille de leur classe par 2 reviendrait donc à réduire les effectifs de 12 élèves, ce qui devrait avoir un effet de $12 \times 0.95 = 11.4\%$ d’écart-type³³.

En outre, les auteurs trouvent que l’effet d’une réduction de la taille des classes est 2.76 fois plus fort parmi les élèves scolarisés dans des établissements de l’éducation prioritaire que

32. L’intérêt de ce travail est qu’il estime ces effets dans le second degré. Il faut noter que les effets que les auteurs obtiennent dans le premier degré sont du même ordre de grandeur que ceux obtenus dans l’expérience contrôlée américaine STAR (Krueger, 1999), ce qui rend leur méthode d’estimation très crédible. En utilisant une expérience non-randomisée menée en France au début des années 2000 au niveau CP, Bressoux et Lima (2011) obtiennent des résultats également comparables.

33. Dans Piketty et Valdenaire (2006), l’effet d’une diminution du nombre d’élèves par classe est identifié à partir de règles qui conduisent à l’ouverture d’une nouvelle classe quand le nombre d’élèves dans un niveau scolaire est supérieur à un seuil. Lorsque l’on passe de 1 à 2 classes pour un niveau donné, cela conduit en effet à diviser par deux le nombre d’élèves par classe. Multiplier par 12 l’effet d’une réduction de la taille des classes d’un élève ne repose donc pas sur une extrapolation hasardeuse, puisque dans l’étude de Piketty et Valdenaire (2006), cet effet est identifié à partir de réductions des effectifs du même ordre.

dans le reste de la population (cf. Piketty et Valdenaire, 2006, Tableau 12). Diviser la taille des classes par 2 dans les établissements de l'éducation prioritaire devrait donc produire un effet de l'ordre de $11.4 \times 2.76 = 31.5\%$ d'écart-type.

L'effet produit par Sourdun (20.5% d'écart-type) est donc plus fort que celui que l'on serait en droit d'attendre si l'on divisait par 2 les effectifs des classes de collège et lycée en population générale (11.4%), mais moins fort que celui que l'on obtiendrait si l'on menait cette même politique seulement dans les établissements de l'éducation prioritaire (31.5%). La population scolarisée à Sourdun est socialement défavorisée, ce qui la rapproche sans doute de la population scolarisée dans les établissements de l'éducation prioritaire. Mais son niveau scolaire initial la rapproche de la moyenne des élèves français. Au total, et si l'on tient compte du fait qu'il existe une imprécision statistique dans toutes ces estimations, on peut sans doute considérer que l'effet de Sourdun est à peu près comparable à ce que l'on obtiendrait si l'on divisait la taille des classes par deux au sein d'une population similaire, ce qui constitue une politique alternative d'un coût semblable³⁴.

6.1.5 L'effet de Sourdun n'est pas significativement différent parmi les garçons et les filles, les collégiens et les lycéens, les élèves forts et les élèves faibles.

La Table 18 présente l'effet de Sourdun par sous-groupe. Nous présentons d'abord l'effet de Sourdun en distinguant les garçons des filles, puis en distinguant les collégiens des lycéens, et enfin en distinguant les élèves initialement forts des élèves initialement faibles. A l'issue de chacune de ces analyses, nous conduisons un test statistique pour déterminer si l'effet de Sourdun est significativement plus fort dans l'un des deux sous-groupes, et nous reportons la statistique résultant de ce test (*p-value*)³⁵.

En français, aucun effet par sous-groupe n'est significativement différent de 0. En mathématiques, l'effet à deux ans chez les filles n'est pas significativement différent de l'effet chez les garçons. Quand on compare l'effet chez les collégiens et les lycéens on ne trouve pas de différence significative non plus. Enfin, l'effet est quasiment le même chez les élèves initialement faibles que parmi les élèves initialement forts. Au total, on ne peut pas affirmer que l'effet de Sourdun est plus fort dans un sous-groupe que dans un autre. C'est peut-être dû en partie à la petite taille de notre échantillon qui ne nous permet pas de détecter des différences fines entre les groupes. Il reste qu'on ne peut rejeter, avec les données disponibles, l'hypothèse selon

34. Dans cette analyse coût bénéfice, nous avons pris le parti de négliger les effets non significativement différents de 0. Les prendre en compte ne change pas radicalement le constat, puisque l'on trouve toujours que l'effet de Sourdun est plus fort que l'effet d'une division par 2 de la taille des classes en population générale, et plus faible que l'effet de cette politique en ZEP.

35. La *p-value* d'un test statistique est le niveau minimal auquel on peut rejeter l'hypothèse testée. Plus une *p-value* est faible, plus on peut être sûr que l'hypothèse doit être rejetée. En général, on considère que l'on peut rejeter une hypothèse si la *p-value* du test correspondant est inférieure à 0.05.

laquelle l'internat a produit ses effets sur tous les types d'élèves. Cette absence de différences permet au moins de tirer une conclusion intéressante. Le départ en internat semble tout aussi productif sur les collégiens que sur les lycéens, alors qu'on aurait pu craindre que des élèves plus jeunes soient davantage pénalisés d'être séparés de leurs familles.

TABLE 18 – Effets sur les scores cognitifs par sous-groupe

Matière	Sous-groupe	Moyenne Témoin	Effet d'une 1 ^{ère} année à Sourdun	σ	Effet d'une 2 ^{ème} année à Sourdun	σ	N
Français	Filles	0.087	0.130	0.158	0.000	0.180	436
	Garçons	-0.087	-0.387	0.281	-0.213	0.223	308
	<i>P-value</i>		<i>0.102</i>		<i>0.301</i>		
	Collégiens	0.009	-0.200	0.223	-0.171	0.192	355
	Lycéens	0.051	0.061	0.190	-0.012	0.212	389
	<i>P-value</i>		<i>0.361</i>		<i>0.437</i>		
Mathématiques	Filles	0.009	-0.039	0.162	0.460***	0.170	430
	Garçons	0.032	0.060	0.217	0.300	0.235	305
	<i>P-value</i>		<i>0.719</i>		<i>0.425</i>		
	Collégiens	-0.016	-0.133	0.214	0.551***	0.212	352
	Lycéens	0.044	0.108	0.152	0.280	0.180	383
	<i>P-value</i>		<i>0.361</i>		<i>0.142</i>		
	Filles	0.481	-0.045	0.199	0.395**	0.187	382
	Garçons	-0.452	0.032	0.158	0.419**	0.206	350
	<i>P-value</i>		<i>0.761</i>		<i>0.899</i>		

6.1.6 Impact sur les résultats au brevet.

Pour mesurer l'impact de l'internat sur les compétences cognitives des élèves, nous disposons également des résultats (décision et notes) au diplôme national du brevet pour les élèves ayant passé cet examen dans les deux ans suivant leur candidature à Sourdun. Nous pouvons donc mesurer l'impact d'être scolarisé à Sourdun sur l'obtention du brevet au sein de ce sous-groupe d'élèves. Il faut cependant souligner d'emblée une limite forte de cette analyse. Le résultat au brevet dépend en partie des notes obtenues dans le cadre du contrôle continu en troisième. Si Sourdun et les établissements fréquentés par les élèves témoins ont des critères de notation ou des niveaux d'exigence différents, la comparaison des résultats au brevet de ces deux groupes pourra aussi bien refléter des différences de compétences réelles entre les élèves que des différences de notation entre Sourdun et les autres établissements.

Le premier panel de la table 19 présente l'effet de la scolarité à Sourdun sur les résultats au brevet. Nous disposons de la décision d'attribution du diplôme et de la mention éventuelle pour environ 80% des collégiens ayant passé cet examen après leur candidature à l'internat, et cette proportion n'est pas statistiquement différente dans les deux groupes. Tous les internes ont obtenu le brevet, contre seulement 78.3% des candidats dans le groupe témoin. Ils sont également plus nombreux à l'avoir obtenu avec une mention. Les effets estimés sur l'obtention des différentes mentions sont tous positifs, mais seul l'effet sur l'obtention d'une mention assez bien est statistiquement différent de zéro. Sourdun a donc un effet positif sur l'obtention du brevet, mais nous ne pouvons pas dire si cela reflète un écart de compétences produit par l'internat, des pratiques de notation différentes à Sourdun, ou simplement le fait que les élèves à l'internat sont moins nombreux à ne pas s'être présentés aux épreuves terminales.

Si l'on voulait utiliser le brevet pour mesurer l'impact de Sourdun sur les compétences des élèves, il faudrait comparer leurs notes aux épreuves terminales, puisqu'elles ne sont pas sujettes au biais de notation mentionné ci-dessus³⁶. Néanmoins, ces données ne sont disponibles que lorsque les responsables légaux des élèves ont autorisé leur divulgation. Or ces notes manquent pour 54% des élèves témoins, contre seulement 16% parmi les élèves de Sourdun (table 19, second panel). Cela risquerait fort de biaiser la comparaison des deux groupes, puisque nous devrions comparer les notes de la grande majorité des élèves de Sourdun à celles d'un groupe témoin amputé de 54% de ses élèves. Les élèves dont les parents ont refusé de divulguer leurs notes sont aussi plus nombreux à avoir échoué au brevet, et moins nombreux à avoir obtenu une mention "Très bien"; il ne s'agit donc pas d'élèves représentatifs du reste de la population. Les deux échantillons ne sont ainsi pas comparables et les résultats sur les notes de l'examen terminal au brevet ne peuvent pas être utilisées pour mesurer rigoureusement l'impact de Sourdun sur les compétences cognitives des élèves.

6.2 Impact sur les aspirations, la motivation et la confiance en soi des élèves

Comme de nombreux travaux en sociologie de l'éducation l'ont montré (voir par exemple Boudon, 1973), l'accès aux études supérieures n'est pas qu'une affaire de compétences et de niveau scolaire. L'information dont disposent les élèves et leurs familles, les représentations qu'ils se font du champ des possibles, leur confiance en eux, ainsi que le plaisir qu'ils prennent à leurs études sont autant de facteurs qui guideront leur choix de s'orienter ou non vers des études supérieures, et la filière qu'ils choisiront. Tous ces déterminants dits non-cognitifs

36. En effet, ces notes ne sont pas attribuées par les professeurs des élèves. Nous n'avons pas de raison de penser que les professeurs qui ont corrigé les épreuves de brevet des élèves de Sourdun ont des pratiques de notation très différentes de ceux qui ont corrigé les épreuves des élèves témoins.

TABLE 19 – Effet sur le brevet

	Moyenne Témoin	Sourdun - Témoin	σ	N
Décision et mention				
Résultat au brevet manquant	0.193	-0.046	0.109	129
A obtenu le brevet	0.783	0.218**	0.098	111
Mention Très bien	0.083	0.012	0.123	111
Mention Bien	0.171	0.119	0.159	111
Mention Assez bien	0.178	0.338*	0.176	111
Reçu sans mention	0.351	-0.250	0.168	111
A échoué au brevet	0.064	-0.090	0.067	111
Absent aux épreuves terminales	0.082	-0.128	0.085	111
Epreuves terminales				
Notes manquantes	0.541	-0.38*	0.229	129

expliquent sans doute une part importante des inégalités d'accès aux filières plus sélectives que l'on observe entre les élèves issus des milieux populaires et les autres. Parfois, ces élèves ne savent tout simplement pas que ces filières existent ; lorsqu'ils en ont entendu parler, elles leur semblent trop souvent inaccessibles, faute d'avoir dans leur entourage des proches ayant suivi ce type de formation ; enfin, ces options peuvent simplement ne pas leur sembler attractives, parce que les études ne sont parfois pas autant valorisées dans leur milieu social que dans d'autres.

D'autre part, si de nombreuses études montrent le rendement salarial des diplômés et des compétences cognitives sur le marché du travail, un ensemble de travaux récents met en avant le rôle comparable de facteurs non cognitifs tels que la confiance en soi et les capacités relationnelles (voir par exemple Heckman, Malofeeva, Pinto, et Savelyev, à paraître).

Il est donc possible que les effets de Sourdun sur d'autres dimensions que le niveau scolaire des élèves soient tout aussi déterminants pour l'avenir des jeunes qui y sont scolarisés. En s'appuyant sur les échelles psychométriques disponibles, nous avons cherché à mesurer l'impact de l'internat d'excellence de Sourdun sur les aspirations des élèves, sur leur motivation et leur goût pour les études, et sur l'image qu'ils se font d'eux-mêmes et de leur rapport aux autres.

La proportion d'élèves qui souhaitent faire des études supérieures est à peu près la même à Sourdun et dans le groupe témoin (environ 80%). Néanmoins, les élèves à Sourdun souhaitent entreprendre des études plus sélectives : 27.4% d'entre eux veulent intégrer une classe préparatoire contre 9.7% dans le groupe témoin, et 71.4% veulent obtenir au moins un Master contre 57.1% (premier panel de la table 20)³⁷. En revanche, on n'observe pas d'effet sur la

37. En ce qui concerne les classes préparatoires, cette différence vient peut-être du fait qu'il en existe une à Sourdun, si bien que les élèves y sont plus familiers avec cette voie que ceux du groupe témoin.

confiance que les élèves ont en leur capacité d'atteindre ces objectifs : si tous les effets estimés sont positifs, aucun n'est significativement différent de 0 (second panel de la table 20).

TABLE 20 – Études supérieures : ambition et confiance en soi

	Moyenne Témoin	Sourdun - Témoin	σ	N
Ambition : l'élève souhaite...				
Faire des études supérieures	0.800	0.038	0.060	362
Intégrer une classe préparatoire	0.097	0.177***	0.064	362
Obtenir au moins un master	0.571	0.143*	0.082	362
Confiance en soi : l'élève se sent capable...				
D'obtenir le baccalauréat	0.886	0.004	0.047	361
D'intégrer une classe préparatoire	0.409	0.055	0.092	360
D'obtenir un master	0.518	0.145	0.090	360

La mesure de la motivation a fait l'objet de nombreuses recherches en psychologie. Les psychologues distinguent trois concepts liés à la motivation (Deci et Ryan, 1985). On parle de motivation intrinsèque lorsqu'une personne entreprend une activité pour le plaisir et la satisfaction qu'elle en retire. On parle au contraire de motivation extrinsèque lorsqu'une activité est entreprise pour des raisons instrumentales, pour en retirer quelque chose de plaisant, ou pour éviter quelque chose de déplaisant. Enfin, l'amotivation correspond simplement à l'absence de toute forme de motivation. Pour mesurer la motivation des élèves vis-à-vis de leur scolarité, nous avons utilisé trois échelles psychométriques conçues et validées par Vallerand, Blais, Brière, et Pelletier (1989). L'échelle de motivation intrinsèque est construite en demandant aux élèves s'ils sont d'accord avec des affirmations telles que "Je vais au collège/lycée pour le plaisir de découvrir de nouvelles choses jamais vues auparavant." Plus un élève déclarera être en accord avec ce type d'affirmation, plus il obtiendra un score élevé sur l'échelle de motivation intrinsèque. Pour construire l'échelle de motivation extrinsèque, on demande aux élèves s'ils sont d'accord avec des affirmations telles que "Je vais au collège/lycée pour pouvoir décrocher un meilleur emploi plus tard." Enfin, pour construire l'échelle d'amotivation, on leur demande d'indiquer s'ils sont d'accord avec des affirmations telles que "Honnêtement je ne sais pas pourquoi je vais au collège/lycée ; j'ai vraiment l'impression d'y perdre mon temps."

Le fait d'être scolarisé à l'internat d'excellence de Sourdun a un fort effet positif sur la motivation scolaire intrinsèque, mais n'a pas d'effet sur la motivation extrinsèque et sur l'amotivation. La table 21 montre en effet que sur les échelles de motivation extrinsèque et d'amotivation, les scores des élèves à Sourdun ne diffèrent pas significativement de ceux des élèves témoins. En revanche, sur l'échelle de motivation intrinsèque, le score des élèves à Sourdun est significativement plus élevé que celui des élèves témoins. La différence est

de l'ordre de 50% d'un écart-type. Nous manquons malheureusement de points de référence pour mettre ces effets en perspective, car les études qui évaluent des politiques éducatives considèrent en général seulement leurs effets sur le niveau scolaire des élèves, pas sur leur motivation.

TABLE 21 – Motivation scolaire, 2 ans après la candidature

	Moyenne Témoïn	Sourdun - Témoïn	σ	N
Motivation externe	-0.022	-0.034	0.183	362
Amotivation	0.092	-0.299	0.208	362
Motivation interne	-0.054	0.501***	0.181	362

Enfin, nous avons mesuré l'effet de Sourdun sur l'estime que les élèves ont d'eux-mêmes, et sur l'image qu'ils se font de leur rapport aux autres. Ici, comme dans le tableau 17 discuté plus haut, nous avons utilisé la traduction française du *Self Perception Profile for Adolescents* (Bouffard, Seidah, McIntyre, Boivin, Vezeau, et Cantin, 2002). Cette échelle psychométrique mesure l'estime que les élèves ont d'eux-mêmes au moyen de différentes questions. Dans le tableau 17, on examinait l'estime de soi scolaire des élèves. On analyse ici d'autres dimensions : "estime de soi générale" et "estime de soi sociale". Pour mesurer l'estime de soi générale, on demande par exemple aux élèves de dire quelle affirmation leur correspond le mieux : "Certains jeunes aiment le genre de personne qu'ils sont" ou "D'autres jeunes aimeraient être quelqu'un d'autre"? Pour mesurer l'estime de soi sociale, on leur demande de choisir entre : "Certains jeunes se sentent acceptés par les autres de leur âge", et "D'autres jeunes souhaiteraient que plus de personnes de leur âge les acceptent". Le tableau 22 présente l'impact de la scolarité à Sourdun sur ces deux mesures au bout de deux ans. Alors que Sourdun a des effets positifs sur l'estime de soi scolaire, il n'a d'impact ni sur l'estime de soi générale, ni sur l'estime de soi sociale des élèves.

TABLE 22 – Estime de soi, 2 ans après la candidature

	Moyenne Témoïn	Sourdun - Témoïn	σ	N
Estime de soi sociale	-0.052	0.042	0.184	362
Estime de soi générale	-0.034	0.130	0.204	362

En résumé, la scolarité à l'internat modifie donc les représentations que les élèves se font du champ des possibles, et a aussi un effet positif sur l'intérêt qu'ils prennent à leurs études. On peut donc espérer que l'internat change en profondeur leurs trajectoires scolaires et leur insertion dans l'enseignement supérieur. En revanche, on n'observe pas d'effet sur l'estime que

les internes ont d’eux-mêmes, ni sur l’image qu’ils se font de leurs relations aux autres.

7 L’impact de la scolarité à Sourdu sur les familles des élèves

La question des effets induits de l’internat d’excellence sur d’autres publics que ses bénéficiaires directs³⁸ est fondamentale pour porter une appréciation globale sur le dispositif. Les effets positifs sur les internes s’accompagnent-ils d’effets d’entraînement sur leurs frères et sœurs et leurs parents ?

Le dispositif d’évaluation permet de quantifier de tels effets en étendant les comparaisons test – témoin aux parents et à la fratrie des élèves des deux groupes. Cependant, il faut être attentif au fait que les effets attendus, dans la mesure où ils sont indirects, sont vraisemblablement plus faibles que les effets directs analysés jusqu’à présent. L’absence d’effet significatif doit donc être interprétée avec prudence : elle peut masquer des effets modestes que l’outil statistique ne détecte pas.

7.1 Impact sur les parents

Au-delà des effets sur l’élève lui-même, la scolarisation d’un enfant en internat d’excellence est susceptible d’impacter sa famille dans de multiples dimensions. Cette section aborde un certain nombre d’entre elles, à partir de l’information collectée auprès des parents des élèves candidats à Sourdu.

Tout d’abord, le regard des parents sur l’expérience que vit leur enfant à l’école a pu être modifié. La première partie du tableau 23 montre que les parents des élèves de Sourdu ont une vision particulièrement positive des relations élèves-professeurs. Cela corrobore ce que les élèves eux-mêmes déclarent (voir tableau 11). Néanmoins, les parents du groupe témoin ont une perception beaucoup plus défavorable des professeurs que leurs enfants eux-mêmes (par exemple, seulement 16% des parents pensent que les professeurs s’intéressent au bien-être des élèves témoins, alors que c’est l’avis de 58% des élèves concernés). Sur cette base, l’amélioration de la perception des parents provoquée par la scolarisation à Sourdu (un gain de 36 points de pourcentage de perceptions positives) constitue une multiplication par plus de trois ; l’amélioration de l’opinion des élèves (+31 points de pourcentage) est proportionnellement moins importante.

Les parents des élèves de Sourdu ont aussi une vision plus positive des relations élèves-CPE et élèves-assistants d’éducation, ce qui tranche avec les déclarations des élèves (tableau 11). On peut peut-être voir ici une forme de gratitude des parents vis-à-vis de l’équipe de

38. En économie, on parle d’“effets externes” ou “externalités”.

l'internat, puisque c'est désormais elle qui doit veiller à l'accomplissement des devoirs ou maintenir la discipline dans les chambrées.

Le regard des parents sur le comportement de leur enfant pourrait également être affecté par la scolarité à Sourdu. Les résultats présentés dans la dernière partie du tableau 23 ne suggèrent cependant pas de changement majeur : les parents des élèves de Sourdu ne pensent pas que leur enfant est plus soigneux, fume moins ou a des amis qui ont une meilleure influence. Bien sûr, cette absence d'effet peut tout aussi bien signifier que Sourdu n'a pas eu d'impact dans ces domaines, ou que les parents n'ont pas perçu les changements intervenus.

TABLE 23 – Regard des parents sur l'expérience scolaire de l'enfant

	Moyenne Témoïn	Sourdu - Témoïn	σ	N
Relation avec les enseignants, CPE et surveillants				
Les élèves s'entendent bien avec les professeurs	0.160	0.127	0.080	274
Les professeurs s'intéressent au bien-être des élèves	0.162	0.356***	0.085	278
Il sont réellement à l'écoute de l'enfant	0.203	0.280***	0.096	275
Si besoin, ils lui apporteront une aide supplémentaire	0.240	0.229**	0.092	274
Ils traitent l'enfant avec justice	0.208	0.289***	0.094	270
Les élèves s'entendent bien avec les CPE/surveillants	0.250	0.084	0.093	273
Ils s'intéressent au bien-être des élèves	0.188	0.189**	0.085	273
Ils sont réellement à l'écoute de l'enfant	0.169	0.291***	0.083	273
Si besoin, ils lui apporteront une aide supplémentaire	0.226	0.205**	0.094	267
Ils traitent l'enfant avec justice	0.247	0.207**	0.091	270
Comportement et environnement social de l'enfant				
Heure de coucher samedi dernier	22.785	0.137	0.257	287
Bonne influence de ses amis	0.581	0.118	0.094	296
Enfant soigneux	0.473	-0.035	0.088	296
Enfant fume	0.185	-0.031	0.073	296

La direction dans laquelle Sourdu devrait changer l'implication des parents dans la scolarité de leur enfant est *a priori* ambiguë : le fait que l'enfant soit plus loin pendant la semaine est susceptible de réduire l'information et l'implication des parents ; mais la perception d'un contexte scolaire plus favorable ou d'opportunités nouvelles peut au contraire conduire les parents à s'impliquer et s'informer davantage. Le haut du tableau 24 suggère une légère perte d'information qui s'explique assez naturellement par l'éloignement : les parents des élèves de Sourdu sont deux fois moins souvent capables d'évaluer le temps que leur enfant passe à faire ses devoirs pendant la semaine que les parents des élèves témoins ; en revanche, ils sont bien au courant du temps passé sur les devoirs le week-end (et leur observation d'environ 3/4 d'heure de plus que pour les élèves témoins est très proche de la différence de 35-40 minutes qu'on obtient à partir des déclarations des élèves). Par ailleurs, les parents des élèves de Sourdu

parlent aussi fréquemment de l'école avec leur enfant que les parents des élèves non scolarisés à Sourduun.

En revanche, la scolarisation à Sourduun a un impact significatif sur l'ambition scolaire pour l'enfant (bas du tableau 24). Par exemple, la proportion de parents qui souhaitent que leur enfant obtienne un master augmente de plus de 20 points de pourcentage (elle est de 63% dans le groupe témoin). Il est intéressant de constater que ces effets sont comparables aux effets sur l'ambition des élèves eux-mêmes : on a vu ainsi que la proportion d'élèves désirant obtenir un master augmente de 14 points de pourcentage avec la scolarisation à Sourduun, sur une base de 57% dans le groupe témoin (tableau 20).

TABLE 24 – Implication des parents dans la scolarité de l'enfant

	Moyenne Témoin	Sourduun - Témoin	σ	N
Implication dans la vie scolaire				
Parle de l'école avec lui	0.740	0.013	0.075	296
Ignore sa moyenne en français	0.511	0.095	0.079	395
Ignore sa moyenne en mathématiques	0.541	0.023	0.081	395
Heures de devoirs/jour la semaine	1.600	0.301	0.220	194
Ignore son temps de devoirs la semaine	0.171	0.231***	0.082	296
Heures de devoirs/jour le weekend	1.596	0.751***	0.261	273
Ignore son temps de devoirs le weekend	0.018	0.059	0.038	296
Ambition scolaire				
Enfant capable de faire des études	0.787	0.106	0.073	230
Souhaite qu'il fasse des études	0.786	0.179**	0.082	239
Enfant capable d'obtenir un master	0.599	0.186*	0.103	230
Souhaite qu'il obtienne un master	0.634	0.223**	0.102	239

7.2 Impact sur la fratrie

Le tableau 25 décrit les éventuels effets de la scolarisation d'un enfant à Sourduun sur le regard que les parents portent sur la scolarité des autres enfants du foyer. Aucun effet n'apparaît significatif. Il convient certes d'être prudent, dans la mesure où on attend que ces effets indirects soient plus faibles que les effets directs sur l'enfant scolarisé. La capacité de détection de notre dispositif est peut-être simplement trop faible pour repérer d'éventuelles différences. Ainsi, si on fait abstraction de la marge d'erreur statistique élevée, on observe une hausse de 6.7 points de la part de parents qui pensent que leur enfant, frère ou sœur d'un élève de Sourduun, est capable d'obtenir un master. Cet effet n'est pas nul, mais il est trois fois plus faible que l'effet concernant l'enfant à Sourduun.

Un autre constat que l'on peut tirer du tableau 25 est que les frères et sœurs des élèves

de Sourduun ne bénéficient pas plus souvent d’une chambre individuelle pendant la semaine. En résumé, il ne semble pas que la scolarisation d’un frère ou d’une sœur à Sourduun modifie très fortement la situation de sa fratrie.

TABLE 25 – Regard de la famille sur les autres enfants du foyer

	Moyenne Témoiu	Sourduun - Témoiu	σ	N
Implication dans la vie scolaire				
Parle de l’école avec lui	0.652	0.107	0.096	293
Temps de devoirs insuffisant	0.518	0.007	0.091	285
Heures de devoirs/jour la semaine	1.195	0.234	0.220	262
Heures de devoirs/jour le weekend	1.305	0.096	0.413	262
Ambition scolaire				
Enfant capable de faire des études	0.747	0.119	0.095	215
Souhaite qu’il fasse des études	0.937	-0.028	0.034	328
Enfant capable d’obtenir un master	0.525	0.067	0.109	215
Occupation des chambres				
Chambre individuelle la semaine	0.404	-0.001	0.088	392
Chambre individuelle le weekend	0.360	-0.069	0.093	392
Passage en chambre individuelle	0.138	0.007	0.080	381

8 L’impact de Sourduun sur les élèves qui n’en bénéficient pas

Les internats d’excellence conduisent certains bons élèves à quitter leur établissement scolaire. Cela pourrait avoir un impact négatif sur les élèves qui restent dans ces établissements, ceux-ci ne bénéficiant plus de l’influence positive de leurs camarades partis à l’internat. Mais il n’est pas non plus exclu que le fait de voir un de leurs camarades partir à l’internat ait des effets positifs sur les élèves qui restent. Cela pourrait en effet les conduire à travailler plus, afin d’obtenir les notes suffisantes pour bénéficier à leur tour de cette opportunité.

Notre dispositif expérimental nous permet de mesurer l’impact du départ en internat d’un élève sur leurs camarades qui restent. Partant de l’ensemble des camarades de classe des candidats à Sourduun au moment de leur candidature, on peut comparer, après un an, les résultats scolaires des camarades des élèves du groupe test (orientés vers l’internat) aux résultats des camarades des élèves témoins. Le tirage au sort nous assure que ces deux groupes étaient comparables initialement. Si l’on observe une différence, on pourra l’attribuer au fait que les camarades d’élèves orientés ont vu un élève de leur classe partir à l’internat.

Le tableau 26 compare les résultats au brevet des élèves qui, en 4ème, étaient camarades de classe d’un élève orienté vers l’internat, à ceux des élèves qui étaient camarades de classe

d'un élève témoin. Aucune différence dans cette table n'est statistiquement significative. Ces deux groupes n'ont pas obtenu de résultats différents au brevet, que ce soit en termes de taux d'obtention, de taux de mention, ou de notes aux épreuves terminales et au contrôle continu. Le départ d'un de leurs camarades en internat ne semble donc pas avoir eu d'impact sur les résultats scolaires des autres élèves.

Il n'est pas exclu qu'un déploiement de cette politique à plus grande échelle finisse par avoir un impact sur les élèves qui restent dans des établissements traditionnels. Ici, nous sommes seulement capables d'affirmer que le départ d'un élève n'a pas eu d'impact sur les autres. Rien ne dit qu'on observerait toujours ce résultat si ce n'était pas un mais cinq élèves, par exemple, qui étaient partis en internat.

TABLE 26 – Résultats au brevet des anciens camarades des candidats

	Camarades Témoins	Camarades Sourdu - Camarades Témoins	σ	N
Décision et mention				
Résultat du brevet manquant	0.138	0.042	0.050	2219
A obtenu le brevet	0.712	-0.002	0.029	1803
Mention Très bien	0.061	-0.006	0.021	1803
Mention Bien	0.118	-0.017	0.018	1803
Notes aux épreuves terminales				
Notes manquantes	0.458	0.024	0.041	2219
Epreuve de français	-0.025	0.115	0.094	1145
Epreuve de mathématiques	-0.029	-0.061	0.128	1129
Epreuve d'histoire-géographie	-0.022	0.010	0.100	1135
Contrôle continu				
Contrôle continu manquant	0.440	0.027	0.041	2219
Moyenne contrôle continu	-0.015	-0.002	0.092	1169

9 Conclusion

Le protocole d'évaluation rigoureux et le suivi systématique des élèves et de leurs familles mis en œuvre dans ce travail nous permettent de dessiner les nombreuses spécificités de l'internat d'excellence de Sourdu et d'en mesurer les effets sur les élèves. Le bilan est positif. Au prix certes de moyens importants, l'internat propose aux élèves un véritable cadre de travail qui leur a permis, passé un temps d'adaptation, d'atteindre un niveau en mathématiques beaucoup plus élevé que des élèves semblables mais non scolarisés à Sourdu. L'absence d'effet en français n'est pas une véritable surprise : plusieurs autres études évaluant des politiques scolaires trouvent aussi des effets sur les compétences des élèves en mathématiques mais pas

en langue. Sur le plan non-cognitif, les élèves ont une motivation renforcée pour les études et une plus grande ambition scolaire, également exprimée par leurs parents. Il semble donc que les conditions sont réunies pour que ces élèves connaissent une carrière scolaire et une insertion professionnelle plus réussies. Au-delà des mesures de court terme présentées ici, c'est bien à cette aune qu'il faudra juger cette politique. Un suivi de ces jeunes pendant plusieurs années est possible à partir de données administratives d'éducation et d'emploi, et c'est le projet que nous souhaitons maintenant mener pour prolonger ce travail.

Lorsque l'on observe la stagnation du niveau cognitif des plus faibles dans le système français, le taux de décrochage inchangé depuis plus de dix ans, ou l'incapacité de la politique d'éducation prioritaire à produire des effets visibles, on peut avoir le sentiment que les politiques scolaires sont impuissantes. L'expérience des internats d'excellence démontre au contraire qu'une mobilisation réelle sur un projet ciblé et ambitieux peut produire des effets mesurables. La force de l'internat est d'agréger de nombreuses composantes : moyens renforcés, recrutement sur profil des enseignants, sélection des élèves qui s'affranchit de la carte scolaire, engagement des uns et des autres dans un projet pédagogique et, naturellement, régime d'internat, avec le contrôle qu'il permet d'exercer sur le rythme des élèves. Du point de vue de l'analyse, cependant, c'est un inconvénient, car il n'est pas possible de distinguer le rôle de tel ou tel élément. Peut-être sont-ils tous nécessaires, ou peut-être suffirait-il d'en systématiser quelques-uns dans un cadre plus souple ou moins coûteux. Certaines expériences à l'étranger permettent d'émettre des hypothèses, rien en France ne permet d'aller plus loin pour l'instant.

Cependant, cette politique fait un choix, celui de cibler des ressources importantes sur des élèves motivés, de niveau scolaire médian et regroupés entre eux. On ne peut pas exclure que ce profil d'élèves soit celui sur lequel ce type d'effort est véritablement efficace. Peut-être est-il préférable de concentrer les ressources là où elles sont utiles, plutôt que de les diluer comme l'a longtemps fait la politique d'éducation prioritaire. Mais les internats d'excellence laissent entièrement ouverte une question : que faire pour les autres ? En mettant en perspective les effets de Sourduin avec ceux d'une politique de réduction d'effectifs, nous avons rappelé qu'il existe d'autres options. Réduire substantiellement la taille des classes pour tous les élèves en difficulté serait néanmoins une option extrêmement coûteuse. La littérature de recherche internationale permet d'en envisager d'autres. Il reste à les expérimenter et à les évaluer en France.

Références

- ABDULKADIROGLU, A., J. ANGRIST, S. M. DYNARSKI, T. J. KANE, ET P. A. PATHAK (2011) : “Accountability and Flexibility in Public Schools : Evidence from Boston’s Charters And Pilots,” *The Quarterly Journal of Economics*, 126(2), 699–748.
- ANGRIST, J. D., S. M. DYNARSKI, T. J. KANE, P. A. PATHAK, ET C. R. WALTERS (2010) : “Inputs and Impacts in Charter Schools : KIPP Lynn,” *American Economic Review*, 100(2), 239–43.
- BEFFY, M., ET L. DAVEZIES (2012) : “Has the “Ambition Success” educational program achieved its ambition ?,” Document de Travail, mimeo, CREST.
- BÉNABOU, R., F. KRAMARZ, ET C. PROST (2004) : “Zones d’éducation prioritaire : quels moyens pour quels résultats ?,” *Economie et statistique*, (380), 3–30.
- BOUDON, R. (1973) : *L’Inégalité des chances*. Armand Colin.
- BOUFFARD, T., A. SEIDAH, M. MCINTYRE, M. BOIVIN, C. VEZEAU, ET S. CANTIN (2002) : “Mesure de l’estime de soi à l’adolescence : version canadienne-française du Self-Perception Profile for Adolescents de Harter,” *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, 34(3), 158–162.
- BRESSOUX, P., ET L. LIMA (2011) : “La place de l’évaluation dans les politiques éducatives : le cas de la taille des classes à l’école primaire en France,” *Raisons éducatives*, (15), 99–123.
- COUR DES COMPTES (2010) : “L’éducation nationale face à l’objectif de la réussite de tous les élèves,” Document de Travail, Rapport public thématique.
- CUNHA, F., J. J. HECKMAN, L. LOCHNER, ET D. V. MASTEROV (2006) : “Chapter 12 : Interpreting the Evidence on Life Cycle Skill Formation,” vol. 1 of *Handbook of the Economics of Education*, pp. 697 – 812. Elsevier.
- CURRIE, J. (2001) : “Early Childhood Education Programs,” *Journal of Economic Perspectives*, 15(2), 213–238.
- CURTO, V. E., ET R. G. FRYER (2011) : “Estimating the Returns to Urban Boarding Schools : Evidence from SEED,” NBER Working Papers 16746, National Bureau of Economic Research, Inc.
- DECI, E. L., ET R. M. RYAN (1985) : *Intrinsic motivation and self-determination in human behavior*. Springer.

- DELAHAYE, J.-P., ET A. DULOT (2011) : “La mise en place des premiers internats d’excellence,” Rapport 2011-057, Inspection générale de l’éducation nationale et Inspection générale de l’administration de l’éducation nationale et de la recherche.
- DOBBIE, W., ET R. G. FRYER (2011) : “Are High-Quality Schools Enough to Increase Achievement among the Poor? Evidence from the Harlem Children’s Zone,” *American Economic Journal : Applied Economics*, 3(3), 158–87.
- DORANS, N. J., M. POMMERICH, ET P. W. HOLLAND (2007) : *Linking and aligning scores and scales*. Springer.
- DUTOIT, M. (ed.) (2003) : *IRT from SSI*. Scientific Software International.
- GLASMAN, D. (2010) : “Les usages sociaux de l’internat scolaire,” mimeo.
- HAMBLETON, R. K. (1991) : *Fundamentals of Item Response Theory*. SAGE.
- HECKMAN, J., L. MALOFEEVA, R. PINTO, ET P. SAVELYEV (a paraître) : “Understanding the Mechanisms Through Which an Influential Early Childhood Program Boosted Adult Outcomes,” *American Economic Review*.
- HOPKINS, K. D., ET G. H. BRACHT (1975) : “Ten-Year Stability of Verbal and Nonverbal IQ Scores,” *American Educational Research Journal*, 12(4), 469–477.
- IMBENS, G. W., ET J. D. ANGRIST (1994) : “Identification and Estimation of Local Average Treatment Effects,” *Econometrica*, 62(2), pp. 467–475.
- KHERROUBI, M., ET J.-Y. ROCHEX (2002) : “La recherche en éducation et les ZEP en France. 1ère partie : Politique ZEP, objets, postures et orientations de recherche,” *Revue française de pédagogie*, (140), 103–131.
- KNUDSEN, E. I., J. J. HECKMAN, J. L. CAMERON, ET J. P. SHONKOFF (2006) : “Economic, neurobiological, and behavioral perspectives on building America’s future workforce,” *Proceedings of the National Academy of Sciences*, 103(27), 10155–10162.
- KOLEN, M. J., ET R. L. BRENNAN (2004) : *Test Equating, Scaling, and Linking*. Springer.
- KRUEGER, A. (1999) : “Experimental Estimates of Education Production Functions,” *Quarterly Journal of Economics*, 114(2), 497–532.
- LAVEAULT, D., ET J. GRÉGOIRE (2002) : *Introduction aux théories des tests en psychologie et en sciences de l’éducation*. De Boeck.

- PIKETTY, T., ET M. VALDENAIRE (2006) : *L'impact de la taille des classes sur la réussite scolaire dans les écoles, collèges et lycées français. Estimations à partir du panel primaire 1997 et du panel secondaire 1995*, vol. 173 of *Les Dossiers*. Ministère de l'Éducation nationale, de l'Enseignement supérieur et de la Recherche Direction de l'Évaluation et de la prospective.
- RAYOU, P., ET D. GLASMAN (2012) : "Les internats d'excellence : un nouveau défi éducatif?," mimeo, Institut français de l'éducation.
- ROBERT, B. (2008) : "L'internat, une tradition française à réinventer," *Les cahiers de l'éducation*, (79), 1–7.
- STOCKING, M. L., ET F. M. LORD (1983) : "Developing a common metric in item response theory.," *Applied Psychological Measurement*, 7, 201–210.
- VALLERAND, R. J., M. R. BLAIS, N. M. BRIÈRE, ET L. G. PELLETIER (1989) : "Construction et validation de l'échelle de motivation en éducation (EME).," *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, 21(3), 323.
- WILSON, M. (2004) : *Constructing Measures : An Item Response Modeling Approach*. Lawrence Erlbaum Associates Inc.

Annexe : modèle utilisé pour construire les scores des élèves

Cette annexe présente la méthode suivie par la DEPP pour construire les mesures de compétences des élèves en mathématiques et en français. L'objectif de cette méthode est double. Il s'agit d'abord d'éliminer, parmi les items proposés aux élèves, ceux présentant des caractéristiques anormales (item faiblement corrélé aux autres items, item avec un taux de bonnes réponses variant anormalement d'une année à l'autre, item très bien réussi dans un niveau et très mal réussi dans un autre...). Il s'agit aussi d'utiliser le fait que les tests sont chaînés, avec des items communs d'un niveau à l'autre, pour construire une mesure synthétique des compétences qui permette de placer tous les élèves de la 6ème à la 1ère sur une même échelle.

Fondements théoriques

Théorie des tests

La théorie des tests considère le score mesuré d'un individu comme composé d'un score "vrai" et d'une erreur de mesure. Ce cadre d'analyse est bien décrit dans Laveault et Grégoire (2002). L'indicateur le plus pertinent issu de cette théorie pour juger de la qualité d'un item est le R_{bis} , qui mesure la corrélation item-test. S'il est faible, l'item apporte peu d'information sur les compétences des individus, il est non discriminant. Il est communément admis au sein de la communauté scientifique que si son est R_{bis} est inférieur à 0.2, un item est jugé non discriminant.

Si la théorie des tests donne un indicateur permettant de juger de la qualité d'un item, elle ne propose pas de solution aisée pour comparer la performance d'élèves sur un test composé d'items passés par les élèves de plusieurs les niveaux scolaires (items ancres) et d'items spécifiques à chaque niveau. Les modèles de réponse à l'item offrent cette possibilité.

Le modèle de réponse à l'item

Nous nous plaçons dans le cadre d'un modèle de réponse à l'item à deux paramètres avec une fonction de lien logistique (cf. Hambleton (1991) et Wilson (2004)) :

$$P_{ij} = \Pr(Y_{ij} = 1 | \theta_i, a_j, b_j) = \frac{e^{1.7a_j(\theta_i - b_j)}}{1 + e^{1.7a_j(\theta_i - b_j)}}$$

La probabilité P_{ij} que le sujet i réussisse l'item j est une fonction logistique du niveau de compétence θ_i du sujet i , du paramètre de difficulté b_j de l'item j et du paramètre de discrimination a_j de l'item j . Ce que l'on veut mesurer est θ_i . Cette variable n'est définie qu'à une transformation croissante affine près. En général, sa moyenne est normalisée à 0 et son écart-type à 1.

Fonctionnement différentiel

On parle de fonctionnement différentiel d'un item (DIF) si la réponse de l'individu dépend d'une caractéristique autre que la compétence. Pour détecter un fonctionnement différentiel, on peut comparer le taux de bonne réponse à l'item d'une année sur l'autre ou d'un niveau scolaire à l'autre. À compétence égale, la probabilité de réponse à un item ne doit pas varier en fonction de l'année de passation ou en fonction du niveau du test.

Une façon simple de détecter un DIF à partir de l'année de passation du test consiste à présenter un graphique avec les taux de bonne réponse aux items l'année t sur l'axe des abscisses et les taux de bonne réponse l'année $t+1$ sur l'axe des ordonnées. Cette représentation compare la proportion de bonnes réponses d'une année par rapport à l'autre. Elle permet de repérer des items éloignés d'une droite moyenne qui représente l'évolution globale de la difficulté des items.

Pour détecter un DIF à partir du niveau de passation du test, on peut estimer le modèle décrit plus haut sur chaque niveau scolaire séparément, calculer la moyenne des paramètres de difficulté dans chacun des niveaux, puis comparer la difficulté de chaque item (après normalisation par cette difficulté moyenne) d'un niveau à l'autre. Intuitivement, si on trouve que l'item le plus difficile pour les élèves de 5ème était le plus facile pour ceux de 6ème, on peut considérer qu'on a détecté un DIF. En pratique, les items qui présentent une différence de difficulté supérieure à une constante sont éliminés.

Ajustement des métriques

Estimation séparée et réconciliation ex-post

Après avoir éliminé les items présentant un fonctionnement différentiel, on peut estimer le modèle séparément sur chaque niveau scolaire. Les mesures de compétences des élèves sont définies comme les paramètres $\hat{\theta}_i$. On obtient alors une métrique différente pour chaque niveau scolaire. Un ajustement est donc nécessaire pour placer les compétences des élèves des différents niveaux sur la même échelle. Nous envisageons le cas de deux groupes de sujets non équivalents en termes de compétence, qui ont passé des tests différents comportant des items communs (*common-items non-equivalent group design*, Kolen et Brennan, 2004).

Les paramètres des items communs, estimés indépendamment sur deux groupes différents, sont alors reliés par une relation linéaire. Remarquons que dans le cas du modèle à deux paramètres, si on effectue le changement d'échelle $\theta^* = A\theta + B$, alors les nouveaux paramètres d'items seront tels que $b_j^* = Ab_j + B$ et $a_j^* = a_j/A$. De cette manière, la probabilité estimée pour un sujet de réussir un item est inchangée. Il s'agit donc de déterminer les coefficients A et B tels que les paramètres des items communs soient proches, selon qu'ils sont estimés sur un groupe ou sur un autre. Puis, l'ajustement se déduit en appliquant $\theta^* = A\theta + B$.

De nombreuses méthodes ont été proposées pour estimer A et B (voir Kolen et Brennan, 2004). Par exemple, une méthode très simple, dite *mean/mean*, consiste tout simplement à remplacer les a_j et les a_j^* par leurs moyennes respectives \bar{a} et \bar{a}^* pour calculer A , puis de remplacer les b_j et les b_j^* par leurs moyennes respectives \bar{b} et \bar{b}^* pour calculer B .

Stocking et Lord (1983) ont proposé une procédure plus complexe et plus précise pour estimer A et B (voir Kolen et Brennan, 2004). Soient (a_{j1}, b_{j1}) les paramètres de l’item j estimés sur les élèves du premier groupe, (a_{j2}, b_{j2}) les paramètres de l’item j estimés sur les élèves du second groupe et (a_{j2}^*, b_{j2}^*) la valeurs des paramètres du second groupe transformés sur l’échelle du premier. Il s’agit de déterminer les meilleurs A et B tels $b_{j2}^* = Ab_{j2} + B$ et $a_{j2}^* = a_{j2}/A$ et tels que les b_{j2}^* (respectivement les a_{j2}^*) soient proches des b_{j1} (respectivement des a_{j1}).

Pour cela, la procédure consiste à minimiser l’écart entre les deux scores “vrais” d’une estimation à l’autre (i.e. d’un jeu de paramètres à l’autre). Pour trouver A et B , Stocking et Lord (1983) proposent de minimiser F :

$$F = \frac{1}{n} \sum_{i=1}^n (\xi_i - \xi_i^*)^2$$

où ξ_i et ξ_i^* sont les scores “vrais” des sujets i obtenus sur les J items communs : $\xi_i = \sum_{j=1}^J P_{j1}(\theta_i)$ et $\xi_i^* = \sum_{j=1}^J P_{j2}^*(\theta_i)$ avec P_{j1} et P_{j2}^* les probabilités de réussite estimées respectivement avec les paramètres (a_{j1}, b_{j1}) et (a_{j2}^*, b_{j2}^*) .

Cela revient à dire qu’à niveau de compétence θ donné, les scores “vrais” obtenus aux items communs par les deux groupes devraient être sinon égaux, du moins très proches, quel que soit le groupe considéré. Sur ce principe, A et B sont estimés, puis appliqués à la transformation des θ du deuxième groupe sur le premier.

Un programme de résolution d’équations non-linéaires doit être mis en œuvre. Le programme informatique *ST* de l’Université de l’Iowa a été utilisé ici (cf. Kolen et Brennan, 2004).

Estimation séquentielle

Un autre façon de procéder est d’estimer les paramètres des items communs sur un premier niveau scolaire, puis de fixer ces paramètres lors de l’analyse du second niveau. Ces estimations peuvent être réalisées avec le logiciel BILOG-MG (cf. DuToit, 2003), qui offre la possibilité de fixer les paramètres des items lors de l’analyse des données.

Estimation simultanée

On peut enfin estimer les paramètres des items simultanément sur plusieurs niveaux. En effet la condition est de pouvoir réaliser l’estimation en tenant compte du fait que tous les

individus n'ont pas "vus" les mêmes items et de relaxer l'hypothèse d'une population unique de distribution normale de moyenne 0 et d'écart-type 1. L'estimation est donc réalisée en considérant k populations de distribution normale avec une population de référence pour laquelle la moyenne est 0 et l'écart-type est de 1. Il est important de noter que le logiciel BILOG-MG ne fonctionne que pour 6 à 8 populations différentes.

Analyse des données des internats d'excellence

Afin de faciliter la modélisation et d'obtenir la comparabilité des scores entre les années et les niveaux, des scores ont été calculés en utilisant les modèles de réponse à l'item et des procédures permettant de rapprocher les métriques.

Toutefois il est important de souligner certains aspects de la comparabilité des scores entre les niveaux et les années. Dans la plupart des cas le rapprochement de métrique se fait pour assurer la comparabilité de deux tests dont le construit est identique mais dont les items sont partiellement différents. On parle dans ce cas d'*equating* (cf. Dorans, Pommerich, et Holland, 2007). Dans notre cas, nous ne pouvons pas poser l'hypothèse que le construit mesuré est strictement identique en raison de la variété des compétences mises en œuvre de la sixième à la première. On parle dans notre cas de *linking*.

Procédure utilisée

La procédure suivante a été utilisée pour l'analyse des données. Cette procédure a été réalisée une fois pour le français et une fois pour les mathématiques de façon indépendante.

Une analyse des R_{bis} a d'abord permis d'éliminer les items non discriminants de l'analyse. Puis, afin de minimiser les erreurs de *linking*, nous avons choisi de réaliser une estimation simultanée des paramètres sur les 3 années pour les classes de seconde et troisième (soit 6 populations, une par année et par niveau), puisqu'il s'agit des deux niveaux centraux. Nous avons alors supprimé les items présentant des fonctionnements différentiels. Le repérage des DIFs se fait groupe par groupe : chacun couple niveau/année est comparé deux à deux, si dans au moins une comparaison la différence de difficulté entre les deux estimations est supérieure à 1 alors l'item est éliminé. Enfin, les items pour lesquels la probabilité de bonne réponse prédite était très éloignée du taux de bonne réponse effectivement observé ont été éliminés (items présentant un mauvais *fit*). A l'issue de ces étapes de nettoyage, les paramètres des items et des scores des individus ont été estimés.

Pour les niveaux suivants, une procédure similaire a été réalisée. Une estimation simultanée permet d'éliminer les items présentant des DIFs entre le niveau à estimer et le niveau estimé précédemment. Une fois les DIFs éliminés, les paramètres sont estimés en fixant les paramètres

des items communs aux valeurs du niveau précédent. Cette première estimation est utilisée pour supprimer les items présentant une mauvaise adéquation entre les valeurs prédites par le modèle et les valeurs observées. Enfin, toujours en fixant les paramètres des items communs avec le niveau précédent, les scores sont estimés sur les items restants.

Robustesse et vérification de la méthode utilisée

Afin de vérifier la validité de la démarche, le *linking* a également été réalisé par la méthode de Stocking et Lord (1983). Dans ce cas les scores obtenus pour un niveau était ajustés en fonction des paramètres A et B obtenus par le logiciel *ST*. En outre, d'autres niveaux que la seconde et la troisième ont aussi été estimés conjointement. Ces méthodes ont abouti à l'obtention de résultats comparables.

Les résultats de la modélisation ont aussi été systématiquement confrontés aux résultats obtenus par une comparaison graphique des difficultés moyennes entre chaque groupe (niveau/année). Pour tester la cohérence des résultats, on a également comparé les moyennes des scores entre niveaux consécutifs.

Résultats

Les tableaux 28 et 27 présentent les scores moyens obtenus, par niveau et année. Les différences entre les niveaux scolaires sont délicates à interpréter. Tout d'abord, les élèves candidats à l'internat d'excellence de Sourdun peuvent être d'un niveau scolaire différent d'un niveau à l'autre. Par exemple, il n'est pas exclu que pour une raison ou une autre Sourdun ait attiré des élèves bien meilleurs en 6ème qu'en 5ème. En outre, les effectifs sont faibles si bien que les différences que l'on observe peuvent ne pas être statistiquement significatives. Néanmoins, à quelques exceptions près, on observe bien que le score est croissant avec le niveau scolaire.

TABLE 27 – Résultats en mathématiques

Année	Niveau	N	Moyenne	Ecart-type	Minimum	Maximum
2010	2	200	0,88	1,30	-1,76	5,30
	3	74	0,19	0,93	-1,70	2,84
	4	84	-0,07	1,02	-3,72	2,96
2011	1	197	0,76	1,27	-3,24	5,11
	2	336	0,45	1,15	-3,62	4,43
	3	192	-0,10	1,10	-4,03	3,86
	4	94	-0,10	0,77	-1,47	3,04
	5	69	-0,69	0,95	-3,93	2,00
	6	49	-0,26	1,24	-3,67	4,03
2012	1	379	0,73	1,45	-4,11	5,56
	2	121	0,22	1,12	-3,92	4,42
	3	133	0,03	0,87	-1,34	3,20
	4	97	-0,29	0,80	-1,78	2,95
	5	51	-0,14	1,04	-3,03	1,99

TABLE 28 – Résultats en français

Année	Niveau	N	Moyenne	Ecart-type	Minimum	Maximum
2010	2	204	0,62	0,88	-1,83	3,65
	3	74	0,21	0,99	-2,39	2,95
	4	86	-0,03	0,98	-1,53	3,40
2011	1	210	0,65	0,83	-0,95	3,53
	2	347	0,44	0,88	-2,77	3,59
	3	198	-0,17	1,00	-3,37	2,82
	4	100	0,18	0,86	-2,50	2,50
	5	69	0,40	0,76	-1,17	2,42
	6	49	0,30	0,70	-2,49	1,62
2012	1	386	0,62	0,94	-2,77	3,59
	2	122	0,15	0,97	-3,35	3,62
	3	135	0,14	0,97	-1,95	3,69
	4	97	0,02	0,79	-2,29	1,91
	5	52	0,65	0,87	-1,80	2,99