

Ministère de l'éducation nationale, de la jeunesse et de la vie associative
Direction de l'évaluation, de la prospective
et de la performance

Document de travail

Série « Études »

N°2012-E01

Qui décroche ?

Cédric Afsa

Cette étude cherche à identifier les facteurs qui font qu'un jeune sera davantage exposé qu'un autre au risque de quitter le collège ou le lycée sans diplôme autre que le brevet des collèges. Elle s'appuie sur le panel de la DEPP d'élèves entrés en 6^{ème} en 1995.

Les *Documents de la Depp* ne reflètent pas la position de la Depp ou des SSA.
Ils n'engagent que leurs auteurs.

Qui décroche ?

Résumé : *Aujourd'hui, un jeune sur six quitte le système éducatif sans diplôme lui donnant un minimum de chances de s'insérer sur le marché du travail. Il s'agit là du cas de figure le plus fréquemment rencontré du phénomène de décrochage scolaire, sur lequel les pouvoirs publics agissent en mettant en place des dispositifs pour aider ces jeunes déscolarisés à trouver des solutions d'insertion sociale ou professionnelle.*

Cette étude concerne l'autre versant de la lutte contre le décrochage, celui de la prévention. Elle cherche à identifier les facteurs qui font qu'un jeune sera davantage exposé qu'un autre au risque de quitter le collège ou le lycée sans diplôme suffisant. Pour ce faire, elle s'appuie sur le panel de la DEPP qui a suivi, jusqu'à la fin de leur scolarité, une cohorte représentative d'élèves entrés en 6ème en 1995. Cette source de données permet de repérer très précisément les jeunes sortant sans diplôme et de les caractériser grâce aux nombreuses informations individuelles et contextuelles qu'elle contient.

On estime à environ 150 000 le nombre de jeunes qui ont débuté leur scolarité secondaire en 1995 et l'ont terminée sans autre diplôme que le brevet des collèges. A ceux-là s'ajoutent environ 50 000 autres qui, bien qu'ayant obtenu un diplôme de l'enseignement secondaire professionnel, n'ont pas terminé avec succès le cycle de formation dans lequel ils s'étaient engagés à la suite.

Parmi tous les déterminants de la sortie sans diplôme identifiés grâce aux informations du panel, le niveau de l'élève à son entrée au collège est de loin celui qui joue le plus grand rôle. Il expliquerait à lui seul près de la moitié des sorties sans diplôme. Ceci confirme la nécessité d'intervenir très en amont, dès l'enseignement primaire, pour en limiter le nombre.

Le décrochage scolaire est un sujet de préoccupation croissante des pouvoirs publics et plus particulièrement des ministères chargés de l'éducation. Les décrocheurs sont ces élèves qui arrêtent leurs études avant d'avoir terminé avec succès le cycle d'enseignement dans lequel ils s'étaient engagés. Les jeunes quittant le lycée sans autre diplôme que le brevet des collèges, c'est-à-dire sans diplôme qui leur offre de réelles perspectives d'insertion professionnelle, en sont la figure emblématique. De fait, leur insertion sur le marché du travail est difficile, et l'est de plus en plus. Ainsi, en 2010, 45 % des sortants récents peu

ou pas diplômés étaient au chômage, contre 10 % pour les diplômés du supérieur¹. L'écart était de 25 points il y a seulement 5 ans. Ce sont aussi les premiers touchés en période de crise économique : entre 2008 et 2009, leur taux de chômage a augmenté de 14 points, alors que la situation des plus qualifiés s'est beaucoup moins dégradée (+4 points de taux de chômage).

Dans l'objectif de leur offrir une solution d'insertion sociale ou professionnelle et de les accompagner dans leur démarche, les ministères concernés ont mis en place un système interministériel d'échange d'informations (SIEI) qu'utilisent les quelque 380 plates-formes locales réparties sur le territoire national, chargées de renouer le contact avec les jeunes « perdus de vue » et leur proposer des solutions. Le dispositif ainsi mis en place a par nature une finalité réparatrice, au sens où les pouvoirs publics opèrent après que le jeune ait décroché du système éducatif. L'autre moyen d'action est la prévention, qui consiste à intervenir *ex ante*, lorsque le jeune est encore en études. Pour cela, il faut identifier les facteurs déclencheurs du décrochage. C'est l'objet de la présente étude, qui adopte une approche statistique du phénomène en s'appuyant sur le panel 1995 de la DEPP (voir encadré 1).

Encadré 1

Les données

Le panel 1995 fait partie des grands outils d'observation statistique de la Direction de l'évaluation, de la prospective et de la performance (DEPP) du Ministère en charge de l'éducation nationale. Il s'inscrit dans une longue tradition, qui a débuté avec la constitution d'une cohorte d'élèves scolarisés en sixième recrutés en trois vagues aux rentrées 1972, 1973 et 1974, cohorte qui a été suivie pendant 11 ans.

Le panel 1995 en est une actualisation. L'échantillon couvre 17 830 élèves entrés en sixième (ou en section d'enseignement spécialisé ou adapté) dans un établissement public ou privé de France métropolitaine, nés le 17 d'un mois, à l'exception des mois de mars, juillet et octobre pour obtenir un taux de sondage proche de 1/40.

Le recueil d'informations s'est fait sous plusieurs modes. En 1995, un questionnaire a été adressé aux chefs d'établissement, qui ont fourni en retour des éléments sur la situation de l'élève à la rentrée 1995 (classe fréquentée, nombre d'élèves, langue vivante étudiée, ...), sa scolarité à l'école élémentaire, son niveau à l'entrée au collège (i.e. les résultats obtenus par l'élève aux évaluations à l'entrée en sixième), et quelques informations sur sa famille (taille de la fratrie, caractéristiques socio-démographiques et activités professionnelles des parents, ...). En 1998, une enquête par voie postale a été effectuée auprès des familles. Elle comportait trois parties : l'enfant et sa famille (approfondissement des questions posées dans le questionnaire de recrutement avec, en particulier, une question sur le niveau des contraintes financières perçu par la famille), l'enfant et l'école (déroulement de la scolarité, mesuré objectivement et subjectivement par l'opinion des parents) et l'enfant et ses études secondaires (représentations et attentes des parents). Deux autres questionnaires complètent le dispositif : une enquête sur la procédure d'orientation en fin de 3ème et une enquête auprès des jeunes en 2002. Enfin, la

1. Béatrice Le Rhun et Claude Minni, *Insertion des jeunes sur le marché du travail : évolution récente du chômage selon le niveau de diplôme*, Note d'Information de la DEPP n° 12.09, mai 2012.

situation scolaire de l'élève (ses principaux paramètres et les caractéristiques de l'établissement fréquenté) a été actualisée chaque année par croisement de fichiers administratifs d'élèves ou par questionnaire le cas échéant.

Le taux d'attrition, c'est-à-dire la part des élèves qui sont sortis du panel sans que l'on connaisse leur devenir, s'élève à 14,2 %. Le panel a pu donc suivre le parcours scolaire d'environ 15 300 élèves jusqu'à leur sortie du niveau secondaire. Plus de 88 % de leurs parents ont répondu à l'enquête menée auprès des familles en 1998. Au total, la présente étude sur le décrochage scolaire s'appuie sur un échantillon de 13 500 élèves, qui a été redressé en conséquence. Enfin, les scores aux tests en mathématiques et en français, d'une part, le niveau des contraintes financières des ménages, d'autre part, ont été imputés lorsque l'information manquait.

Cette source de données permet, d'une part, de suivre la scolarité d'un échantillon d'élèves entrés en 6ème la même année (en 1995) et de dater avec précision leur sortie de l'enseignement secondaire en connaissant le niveau d'études atteint, d'autre part, de mettre en rapport les trajectoires ainsi décrites avec les caractéristiques socio-démographiques de ces élèves ainsi que leurs contextes scolaire et familial.

1. Les décrocheurs : de qui parle-t-on ?

Il faut au préalable lever l'ambiguïté du terme « décrocheur », auquel le sens commun donne des contours très variés. Certains parlent de ces élèves nommés « décrocheurs passifs »² qui, tout en restant scolarisés, participent de moins en moins activement aux cours. Ils les opposent aux « décrocheurs actifs » qui, eux, sont déscolarisés et ont quitté le système éducatif avec un niveau jugé insuffisant. Le débat porte alors sur le niveau minimal à acquérir pour ne pas être considéré comme décrocheur. Sur cette question, plusieurs points de vue coexistent.

Il est souvent admis que le seuil minimal est le niveau V de l'Éducation nationale, qui correspond au certificat d'aptitude professionnelle (CAP) ou au brevet d'études professionnelles (BEP). Un jeune qui n'a pas obtenu un de ces diplômes est donc considéré comme ayant décroché du système éducatif. Dans ce cas, il est préférable de parler de jeune *sor-tant sans diplôme* (autre que le brevet des collèges)³, pour le distinguer d'un autre type de décrocheur qui s'est, pour ainsi dire, installé récemment dans le paysage.

Il s'agit très exactement des décrocheurs tels que l'entend l'article L. 313-7 du code de l'éducation. Ce sont, nous dit le texte de loi, « *ces anciens élèves ou apprentis qui ne sont plus inscrits dans un cycle de formation et qui n'ont pas atteint un niveau de qualification fixé par voie réglementaire* ». L'article D. 313-59 précise que le niveau correspond à l'obtention soit du baccalauréat général, soit d'un diplôme à finalité professionnelle de niveau

2. D'autres distinctions ont cours, différenciant par exemple les « non accrochés » qui ne s'intéressent pas aux cours dispensés, les « décrocheurs » qui prennent de plus en plus de distance notamment en s'absentant, et les « décrochés » qui ont perdu tout lien avec l'institution scolaire. Voir notamment : Jean-Robert Pitte (sous la dir.), *Orientation pour tous*, édition François Bourin, 2011.

3. Le diplôme national du brevet n'est pas considéré comme un diplôme comme les autres au sens où la poursuite des études n'est pas conditionnée à son obtention.

V ou IV inscrit au répertoire national des certifications professionnelles.

En clair, être décrocheur au sens de la loi, c'est ne pas avoir terminé avec succès le cycle de formation de second cycle du second degré dans lequel le jeune s'était engagé. Ainsi, est considéré comme décrocheur un élève titulaire d'un CAP, qui poursuit ses études pour le compléter par un second CAP ou pour obtenir un baccalauréat professionnel mais qui arrête sa scolarité sans avoir atteint son objectif. Autrement dit, on peut être décrocheur *et* diplômé. C'est même une situation relativement fréquente. En effet, d'après les données du panel 1995, on estime à 23,6 % la part des élèves entrés en sixième en 1995 qui ont décroché, toujours au sens de la loi, du système éducatif. Un (petit) quart de ces 23,6 % de décrocheurs sont sortis avec un CAP ou un BEP en poche (table 1). Les trois-quarts des décrocheurs sont donc sortis sans diplôme (autre que le brevet des collèges). Ils représentent 17,9 % des élèves de la cohorte de 1995.

Table 1. Répartition des décrocheurs selon leur niveau d'études à la sortie

	Effectifs (%)
Sans aucun diplôme	53,4
Avec uniquement le Brevet	22,6
Avec un CAP ou un BEP	24,0
Ensemble	100,0

Lecture : 24,0 % des décrocheurs de la cohorte d'élèves entrés en 6ème en 1995 ont obtenu un CAP ou un BEP.

Source : DEPP – Panel 1995.

Si on applique ces proportions aux quelque 850 000 élèves de métropole et des départements d'outre-mer inscrits en 6ème en 1995, alors environ 200 000 de ces jeunes ont décroché du système éducatif et 150 000 sont sortis sans diplôme du secondaire⁴.

La suite de l'étude s'intéresse plus particulièrement aux jeunes sortis sans diplôme. Il y a deux raisons à ce choix. D'abord, ces jeunes constituent la majorité – les trois-quarts – des décrocheurs au sens de la loi. Ensuite, les diplômés du secondaire (CAP ou BEP) forment une population à part au sein des décrocheurs. En témoigne leur taux de chômage constaté quelques années après leur sortie de l'enseignement secondaire : fin 2010, 18 points séparent les peu ou pas diplômés (45 % de chômeurs) des titulaires d'un CAP ou d'un BEP (taux de chômage de 27 %)⁵, alors que l'écart entre les détenteurs du brevet uniquement et ceux n'ayant strictement aucun diplôme est de 4 points seulement (respectivement 42 % et 46 % de chômeurs).

Une étude a été conduite au début des années 2000⁶ sur une catégorie plus restreinte

4. Dans toute la suite de l'analyse et sauf mention contraire, les « sans diplôme » sont les jeunes sortis du secondaire sans autre diplôme que le brevet des collèges. On parlera aussi de « peu ou pas diplômés ».

5. Béatrice Le Rhun et Claude Minni, op. cité.

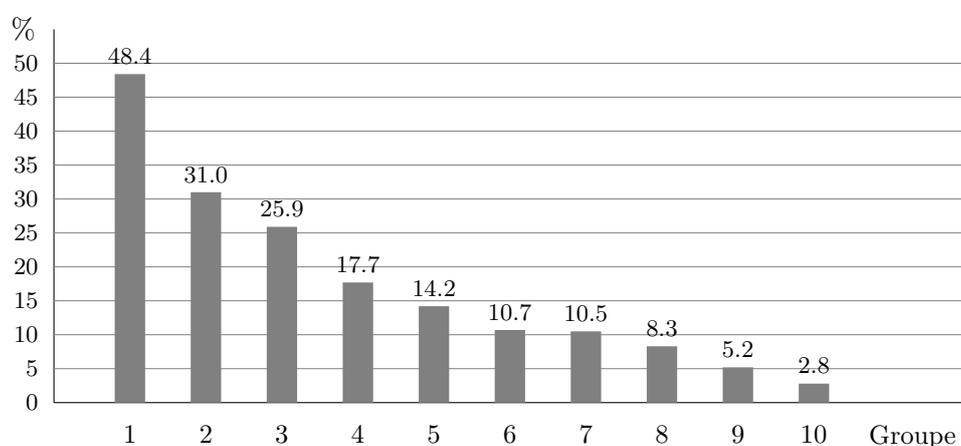
6. Jean-Paul Caille (2000), *Qui sort sans qualification du système éducatif ?* Éducation & formations, DEPP, n° 57.

d'élèves, ceux sortis sans qualification. Il s'agit d'élèves qui ont arrêté leurs études en fin de 3ème ou avant, ou bien, s'ils ont poursuivi en voie professionnelle, n'ont pas atteint l'année terminale de préparation d'un CAP ou d'un BEP. L'article, reposant sur une exploitation du panel 1989 de la DEPP⁷, insistait sur la grande hétérogénéité de cette population en termes de fluidité des parcours antérieurs, de types d'établissements fréquentés (enseignement adapté ou non), de niveau de sortie. L'étude d'Isabelle Robert-Bobée (2012)⁸ confirme la grande diversité des trajectoires des décrocheurs. Elle complète sur ce point la présente analyse, dont l'objectif est avant tout de mettre en évidence les facteurs communs aux jeunes sortis sans diplôme qui les distinguent fondamentalement des diplômés, puis d'identifier les facteurs jouant le plus grand rôle.

2. Les sortants sans diplôme : une population très particulière ?

Un premier constat s'impose : le risque pour un élève de sortir sans diplôme dépend très nettement de son niveau au début du collège. En effet, si on répartit les élèves en dix groupes de niveau selon le score qu'ils ont obtenu aux épreuves de mathématiques à l'entrée en 6ème, la proportion de sortis sans diplôme s'élève à 48,4% pour le premier groupe qui rassemble les 10% d'élèves les plus faibles et à 2,8% pour les 10% d'élèves les plus forts (graphique 1).

Figure 1. Proportion de sortis sans diplôme selon leur groupe de niveau en mathématiques à l'entrée en 6ème



Lecture : Dans le groupe 1, qui rassemble les 10% d'élèves ayant eu les scores le plus faibles aux tests de mathématiques, 48,4% sont sortis sans diplôme.

Source : DEPP – Panel 1995.

D'autres différences existent. Par exemple, les jeunes sortis sans diplôme se recrutent

7. Il s'agit de la version précédente du panel 1995. Il a suivi une cohorte de 27 000 élèves entrés en 6ème en 1989.

8. Isabelle Robert-Bobée (2012), *Les décrocheurs du système éducatif : une diversité de profils*, à paraître.

davantage dans les milieux modestes. A ce titre, les enfants d'ouvriers sont en proportion 4 fois plus nombreux que les enfants de cadres (table 2). Mais cette différence ne tient probablement pas à la seule catégorie sociale. En effet, le score moyen aux tests de mathématiques des enfants de cadres est nettement au-dessus (près de 13 points) de ce qu'ont obtenu les enfants d'ouvriers (table 2). En conséquence, le fait que les enfants d'ouvriers soient plus nombreux à sortir sans diplôme est (peut-être) davantage une affaire de niveau initial que de catégorie sociale.

Table 2. Milieu social de l'élève (cadres *vs* ouvriers) et sortie sans diplôme

Milieu social	Taux de sortie sans diplôme (%)	Score en mathématiques
Cadre (ou prof. int. sup.)	6,3	59,0
Ouvrier	24,9	46,3

Lecture : 24,9% des enfants d'ouvriers, entré(e)s en 6ème en 1995, sont sorti(e)s de l'enseignement scolaire sans diplôme (autre que le brevet), contre 6,3% des enfants – soit près de 4 fois moins – de cadres. Ils ont obtenu un score moyen de 46,3 aux tests de mathématiques.

Source : DEPP – Panel 1995.

La même conclusion s'applique lorsqu'on compare la situation des familles immigrées⁹. Leurs enfants sortent beaucoup plus fréquemment sans diplôme, mais leur niveau en 6ème était nettement plus faible (table 3).

Table 3. Immigration et sortie sans diplôme

Origine	Taux de sortie sans diplôme (%)	Score en mathématiques
Famille immigrée	29,7	43,8
Aucun parent immigré	16,3	52,1

Lecture : 29,7 % des enfants de familles immigrées entrés en 6ème en 1995 sont sortis sans diplôme, à comparer aux 16,3 % des familles dont aucun des parents n'est immigré ; par ailleurs, d'une manière générale, les enfants de familles immigrées ont obtenu un score moyen de 43,8 aux tests de mathématiques qu'ils ont passés à leur entrée en 6ème.

Source : DEPP – Panel 1995.

Ces premiers éléments laissent penser que les sortants sans diplôme ne se distinguent pas aussi nettement des autres qu'on l'envisageait a priori, exception faite de leur niveau scolaire en 6ème, qui serait ainsi *le* facteur clivant. Les liens constatés entre la sortie sans

9. Les familles immigrées sont les ménages où les deux parents dans le cas d'un couple ou le parent seul dans le cas de monoparentalité sont nés étrangers à l'étranger. Une famille biparentale dont un seul des parents est immigré est appelée famille mixte.

diplôme et le milieu social (table 2) ou l'origine des parents (table 3) s'expliqueraient alors avant tout par une question de niveau scolaire au début des études secondaires.

Pour le savoir, il faut recourir à un modèle statistique qui permette de construire une situation (fictive) dans laquelle les sortants sans diplôme auraient débuté leur scolarité secondaire avec le même niveau scolaire que les autres. Ayant ainsi neutralisé la différence de niveau, on examinerait ensuite le rôle résiduel des variables d'intérêt (catégorie sociale, rapport à la migration, ...) sur le processus de sortie sans diplôme.

3. A la recherche de facteurs distinctifs

Approfondir l'analyse consiste à compléter la liste des premières variables que l'on vient d'examiner par d'autres susceptibles elles aussi d'expliquer les différences de profil entre les sortants sans diplôme et les autres, à condition qu'elles soient disponibles dans la source de données. Les variables finalement retenues sont de trois types, plus ou moins fournis en fonction de l'information contenue dans le panel 1995 :

- caractéristiques individuelles des élèves : sexe, scores obtenus aux tests en mathématiques et en français passés en 6ème ;
- environnement familial : structure du ménage (parent isolé/couple), catégorie sociale du ménage (i.e. du chef de famille), diplôme de la mère, présence d'un enseignant dans le foyer, rapport à la migration (en distinguant les familles immigrées et les familles mixtes : voir la note au bas de la page 6), aisance financière du ménage ;
- contexte de l'établissement : zone rurale ou urbaine, taux de chômage local – i.e. de la zone d'emploi – au 2ème trimestre 2000 (au moment où la majorité des élèves suivis par le panel terminent leur seconde).

L'annexe 1 donne les effectifs de l'échantillon selon ces différentes caractéristiques.

Certaines des variables ont été examinées précédemment. La table 2, par exemple, montre un écart de 18,6 points de pourcentage entre les taux de sortie sans diplôme des enfants de cadres et ceux des enfants d'ouvriers. Ces constats se lisent aussi de la manière suivante : en moyenne, la probabilité qu'un enfant d'ouvrier sorte sans diplôme est de 24,9 % ; pour un enfant de cadre, cette probabilité s'établit en moyenne à 6,3 %. Autre exemple, la probabilité pour un élève du secondaire de sortir sans diplôme est en moyenne plus élevée chez les familles issues de l'immigration (29,7 %) que chez les non immigrées (table 3).

Cette manière d'interpréter les tables 2 ou 3 donne la ligne générale à suivre et la nature du modèle statistique à mettre en œuvre pour analyser finement la population des sans diplôme. L'analyse va se centrer sur le lien existant entre un ensemble de caractéristiques socio-démographiques – représentées en l'espèce par les variables listées ci-dessus – et la probabilité qu'un élève a de terminer sa scolarité sans être diplômé. Disant cela, on est déjà dans une démarche de modélisation puisqu'on n'observe pas dans la réalité la probabilité de sortie attachée à chaque élève. On sait seulement si l'élève est ou non diplômé à la fin de ses études secondaires. En d'autres termes, la probabilité individuelle est un construit du modèle. Elle mesure le risque qu'un élève, compte tenu de ses caractéristiques, sorte sans diplôme.

Encadré 2

Le modèle statistique utilisé

Le modèle part de l'hypothèse que la probabilité qu'un élève i de caractéristiques x_i sorte sans diplôme est une fonction des variables x_i , plus précisément d'une combinaison linéaire des x_i . Formellement, cela s'écrit :

$$P(y_i = 1|x_i) = F(x_i\beta)$$

où la variable y vaut 1 si l'élève est sorti sans diplôme et 0 sinon. β représente les coefficients de la combinaison linéaire des variables x . Ce sont les paramètres du modèle. Leurs valeurs sont déterminées en faisant en sorte que le modèle reproduise le mieux possible la réalité.

Il y a plusieurs choix possibles de F . C'est la fonction *logistique* $F(u) = 1/(1 + e^{-u})$ qui a été retenue ici. Le modèle, dénommé *logit*, s'écrit donc :

$$P(y_i = 1|x_i) = \frac{1}{1 + e^{-x_i\beta}}$$

Le choix de la forme fonctionnelle repose sur des considérations théoriques. D'abord, la quantité $1/(1 + e^{-x_i\beta})$ variant entre 0 et 1, $P(y_i = 1|x_i)$ est bien une probabilité. Ensuite, on démontre que, sous des conditions peu restrictives sur les variables x , le modèle *logit* est un modèle d'analyse discriminante, méthode dont un des objets est de répondre à la question suivante : quelles caractéristiques distinguent – discriminent – les deux populations que sont les sortis sans diplôme d'un côté, et les sortis avec diplôme de l'autre¹⁰ ? Enfin, le fait que la probabilité dépende d'une *combinaison linéaire* des variables permet précisément d'évaluer le rôle de chacune d'elles indépendamment des autres. Pour ce faire, on maintient constantes toutes les variables x sauf une dont on fait varier « artificiellement » les valeurs, ce qui permet de mesurer le rôle qu'elle joue elle-même (i.e. les autres variables restant fixées) dans la probabilité P .

Cela étant, cette prise de distance au réel a ses avantages. Le modèle permet, en effet, d'identifier les variables qui différencient réellement les sortis sans diplôme et les élèves diplômés, et d'avoir ainsi une première idée des rôles respectifs joués par différents facteurs a priori déterminants du décrochage scolaire (encadré 2). On saura au moins répondre à des questions restées en suspens. Par exemple : une fois prise en compte la différence de niveau des élèves à l'entrée en 6ème, le milieu social distingue-t-il encore les sortants sans diplôme des autres (table 2) ?

La première lecture des résultats du modèle – table 4 – livre plusieurs enseignements. Elle confirme d'abord le lien précédemment constaté – graphique 1 – entre le niveau de l'élève en 6ème et la sortie sans diplôme. Les paramètres associés aux scores tant en mathématiques qu'en français sont en effet négatifs : plus les scores aux tests sont élevés, moins la probabilité de sortir sans diplôme est importante. Par ailleurs, à même niveau initial en début de collège les enfants d'ouvriers sortent plus souvent sans diplôme que les enfants de cadres : le paramètre associé à la catégorie « ouvriers » est moins

10. Pour de plus amples précisions, voir : C. Afsa, *Le modèle logit dichotomique*, Document de travail de la DEPP, série Méthodes, n° 2011-M02.

Table 4. Facteurs distinctifs des sans diplôme et des diplômés
(modèle logistique)

	Paramètre estimé	Écart-type
Constante	3,144***	0,225
Sexe de l'élève (<i>ref=garçon</i>)		
<i>filles</i>	-0,400***	0,052
Score aux tests de mathématiques (en 6ème)	-0,038***	0,003
Score aux tests de français (en 6ème)	-0,032***	0,003
Structure du ménage (<i>ref=parent isolé</i>)		
<i>parents en couple</i>	-0,459***	0,071
CS du chef de famille (<i>ref=sans activité</i>)		
<i>agric. exploit., artis., commerç., chefs d'entrep.</i>	-0,442***	0,163
<i>cadres, prof. intell. sup.</i>	-0,682***	0,181
<i>prof. intermédiaires</i>	-0,677***	0,162
<i>employés</i>	-0,277*	0,149
<i>ouvriers</i>	-0,353**	0,146
Un parent enseignant ? (<i>ref=non</i>)		
<i>oui</i>	-0,443***	0,160
Diplôme de la mère (<i>ref=non rép., sans objet, aucun ou CEP</i>)		
<i>BEPC, CAP, BEP</i>	-0,461***	0,060
<i>baccalauréat</i>	-0,505***	0,097
<i>diplôme ens. supérieur</i>	-0,674***	0,111
Rapport à la migration (<i>ref=famille non immigrée</i>)		
<i>famille mixte</i>	0,320***	0,110
<i>famille immigrée</i>	0,039	0,076
Le revenu de votre famille est-il ... (<i>ref=très insuffisant ?</i>)		
<i>un peu insuffisant ?</i>	-0,159**	0,067
<i>juste suffisant ?</i>	-0,178***	0,067
<i>tout à fait suffisant ?</i>	-0,313***	0,097
Tranche d'unité urbaine (<i>ref=commune rurale</i>)		
<i>commune urbaine < 5 000 hab</i>	-0,203	0,185
<i>comm. urb. 5 000-10 000 hab</i>	-0,127	0,159
<i>comm. urb. 10 000-20 000 hab</i>	-0,242	0,153
<i>comm. urb. 20 000-50 000 hab</i>	-0,202	0,143
<i>comm. urb. 50 000-100 000 hab</i>	-0,171	0,146
<i>comm. urb. 100 000-200 000 hab</i>	-0,194	0,149
<i>comm. urb. > 200 000 hab</i>	-0,129	0,135
Taux de chômage local (2ème trimestre 2000)	0,011	0,009

Seuils de significativité : *** = 1% ; ** = 5% ; * = 10%.

Lecture : Le paramètre associé à la catégorie sociale « cadres » est négatif. En conséquence, un enfant de cadre, comparativement à un enfant vivant dans une famille dont le chef est sans activité, a une probabilité plus faible de sortir sans diplôme. Par ailleurs, on a moins de 1 % de risques de se tromper en l'affirmant (seuil de significativité à 1 %). Autre exemple : un enfant d'employé a lui aussi moins de risques de sortir sans diplôme qu'un enfant d'une famille dont le chef est inactif (paramètre égal à -0,277), mais l'affirmation est un peu moins certaine puisque le risque d'erreur peut cette fois atteindre 10 % (seuil de significativité à 10 %). Enfin, comme on a plus de 10 % de risques de se tromper en affirmant qu'un enfant d'une famille immigrée risque de sortir sans diplôme davantage qu'un enfant d'une famille non immigrée (paramètre positif égal à 0,039), on préfère ne pas trancher.

Source : DEPP – Panel 1995.

négatif que celui associé à la catégorie « cadres et professions intellectuelles supérieures ». Malgré les marges d'incertitude entourant ces deux valeurs, elles restent différentes¹¹. Le premier constat établi par la table 2 reste valide même en tenant compte entre autres de la différence de niveau initial entre les élèves. En revanche, les enfants de familles immigrées ne se distinguent plus de ceux vivant dans une famille où aucun des parents n'est immigré. Cette fois-ci, la prise en compte simultanée des variables (dont, en particulier, le niveau en mathématiques et en français en 6ème) invalide la première conclusion tirée de la table 3. Plus surprenant est la situation – que nous ne savons pas interpréter – des familles mixtes, c'est-à-dire les couples où un seul des parents est immigré. Les enfants de ces couples sortent apparemment plus souvent sans diplôme¹².

D'autres résultats étaient attendus, comme le rôle protecteur du diplôme de la mère, ou celui du niveau de vie de la famille. On notera aussi l'avantage pour l'élève d'avoir un parent enseignant, ou de ne pas vivre dans une famille monoparentale. La taille de commune ne semble pas jouer de rôle.

Un dernier mot sur la situation du marché du travail, mesurée par le taux de chômage local. Un chômage élevé est susceptible d'avoir sur la trajectoire de l'élève deux effets a priori antagonistes. D'un côté, il peut le pousser à poursuivre ses études et obtenir un diplôme pour avoir de meilleures chances d'insertion sur le marché du travail. Mais il peut aussi dégrader la situation financière de la famille et, au-delà, son climat, et contribuer ainsi à l'arrêt précoce des études. Au vu des résultats, aucun des deux effets ne semble l'emporter.

4. Sortir sans diplôme : pour moitié, une question de niveau en 6ème

A ce stade, les valeurs des paramètres reportées dans la table 4 sont peu parlantes. Elles donnent une information très limitée sur l'importance des rôles joués par les différentes caractéristiques. Pour en avoir une meilleure idée, il faut les transformer en points de probabilité de sortir sans diplôme. Pour chaque modalité de variable, on calcule ce qui est communément appelé l'« effet partiel » ou l'« effet marginal » (voir encadré 3). Chaque effet partiel est un écart « résiduel » au sens où il mesure le rôle joué par la modalité de la variable après avoir annulé, grâce au modèle, les différences tenant aux autres variables.

Encadré 3

Évaluer l'importance des facteurs influant sur la sortie sans diplôme

Il y a plusieurs moyens d'évaluer l'importance du rôle joué par tel ou tel facteur dans le processus de sortie sans diplôme. Le premier est de calculer les *effets partiels* ou *effets marginaux* des différentes variables sur la probabilité de sortir sans diplôme. Second moyen, l'utilisation

11. Un test de comparaison des deux valeurs conclut qu'on a moins de 1% de risques de se tromper en affirmant qu'elles sont différentes, c'est-à-dire que les sans diplôme resteraient proportionnellement plus nombreux chez les enfants d'ouvriers que chez les enfants de cadres, même s'ils partageaient les mêmes caractéristiques observées (niveau en 6ème, diplôme de la mère, revenu de la famille, ...).

12. Le paramètre est positif et statistiquement significatif au seuil de 1% (voir encadré 3).

de critères de qualité du modèle pour estimer l'importance du rôle joué par un ou plusieurs facteurs.

Le calcul des effets partiels (ou effets marginaux)

D'une manière générale, l'effet partiel (ou marginal) d'une variable x_1 sur la probabilité de sortir sans diplôme se mesure par l'amplitude de variation de la probabilité lorsque la variable x_1 varie, les autres variables restant constantes.

Notons d'abord qu'une fois les paramètres β du modèle estimés, on peut calculer pour chaque individu i de l'échantillon sa probabilité, estimée par le modèle, de sortir sans diplôme :

$$\hat{P}(y_i = 1|x_i) = \frac{1}{1 + e^{-x_i\hat{\beta}}}$$

où $\hat{\beta}$ sont les valeurs estimées des paramètres β . Ceci permet alors de calculer l'effet partiel de la variable x_1 . Dans le cas où la variable est binaire (c'est-à-dire vaut 0 ou 1), on procède comme suit. Pour chaque individu i , on calcule la probabilité $\hat{P}(y_i = 1|x_i, x_{1i} = 1)$ en forçant à 1 la valeur de x_1 et en gardant pour les autres variables les valeurs observées pour l'individu i . On calcule ensuite la probabilité $\hat{P}(y_i = 1|x_i, x_{1i} = 0)$ en forçant à 0 la valeur de x_1 et en gardant pour les autres variables les valeurs observées pour l'individu i . Par différence, on obtient l'effet marginal de x_1 sur la probabilité de sortir sans diplôme, pour l'individu i . Pour obtenir l'effet marginal moyen, on fait la moyenne des n différences individuelles, n étant la taille de l'échantillon.

L'extension au cas d'une variable à plus de deux modalités se fait sans problème.

L'utilisation d'un indicateur de qualité du modèle

Pour évaluer l'importance relative des différentes variables, on peut utiliser un des indicateurs mesurant la qualité du modèle. C'est un *critère d'information* qui est retenu ici.

Le critère d'information tire sa dénomination de ce qu'il mesure la perte d'information due au fait qu'on remplace la réalité par un modèle. De ce point de vue, plus la valeur du critère est faible, plus la perte d'information est limitée, et donc meilleur est le modèle au sens où il « trahit » moins la réalité. La démarche est alors la suivante. On estime le modèle avec toutes les variables observées. Puis on le réestime en retirant une variable, mettons x_1 . Le critère d'information va augmenter puisqu'en se privant d'une variable le modèle rend (un peu) moins compte de la réalité. On réestime le modèle en retirant non pas x_1 mais une autre variable, x_2 . Pour les mêmes raisons, le critère d'information augmente. S'il augmente davantage que dans le cas où on retire seulement x_1 , alors on considèrera que la variable x_2 est plus « importante » que x_1 puisque son retrait conduit à une plus grande perte d'information. On peut de cette manière hiérarchiser les apports de différentes variables, ou de différents ensembles de variables. Les critères d'information les plus utilisés sont celui d'Akaike (AIC) et celui de Schwartz (SC).

Un exemple, pour illustrer le propos. On part de l'écart « brut » constaté, tel que reporté dans la table 2, des taux de sortie sans diplôme des enfants de cadres d'une part et des enfants d'ouvriers d'autre part. On se place ensuite dans la situation fictive où les enfants de ces deux catégories auraient le même sexe, auraient eu les mêmes notes aux tests en 6ème, vivraient avec leurs deux parents, etc. Bref, ils ne se différencieraient que sur la catégorie sociale du chef de famille. L'écart « résiduel » est ce que devient l'écart brut une

12. Pour de plus amples précisions, voir : C. Afsa, op. cité.

fois annulées toutes les différences sur les variables autres que la catégorie sociale.

La table 5 donne pour toutes les variables (sauf les trois variables continues : les deux scores aux tests en 6ème et le taux de chômage local) les écarts résiduels (ou effets partiels) et les compare aux écarts « bruts ». Le constat général est que la prise en compte simultanée des facteurs fait substantiellement baisser l'amplitude des écarts. Notamment, l'écart, en termes de taux de sortie sans diplôme, entre les enfants des familles immigrées et ceux des familles non immigrées passe de 13,4 point à une valeur quasi nulle (0,5 point, écart non significatif). On retrouve le résultat déjà mis en avant (voir section 3 *supra*).

Autre exemple, le rôle de la catégorie sociale des parents. On l'a vu, l'écart brut des taux de sortie sans diplôme s'élève à 18,6 points (table 5 : 37,5 - 18,9) entre les enfants de cadres et les enfants d'ouvriers. Lorsque l'on compare la situation des enfants en faisant comme s'il partageaient les mêmes caractéristiques à l'exception de la catégorie sociale de leurs parents, l'écart se réduit à 3,8 points (table 5 : 8,6 - 4,8). En conséquence, 80 %¹³ de l'écart constaté entre les enfants de ces deux catégories sociales s'explique par des différences structurelles (mesurées par les variables prises en compte dans le modèle) entre les deux populations.

Les facteurs observés expliquent donc une très large part des différences constatées *a priori*. Au vu des premières statistiques données en section 2 portant sur le milieu social (table 2) d'une part, et portant sur le rapport à la migration (table 3) d'autre part, on pressent que le niveau de l'élève en 6ème joue un rôle (très) important. Pour en avoir le cœur net, le modèle a été réestimé avec deux variantes. La première a consisté à retenir uniquement les scores aux deux tests passés en 6ème et la variable de catégorie sociale. La seconde variante n'a retenu que les deux scores aux tests en 6ème et la variable « rapport à la migration ». Avec la première variante, l'écart résiduel des taux de sortie sans diplôme entre les enfants de cadres et les enfants d'ouvriers est de 8,1 points. En conséquence, le niveau de l'enfant en sixième (mesuré par les résultats aux deux tests d'évaluation) explique à lui seul plus de la moitié¹⁴ (56 % pour être précis) de l'écart constaté sur ces deux sous-populations. Avec la seconde variante, l'écart résiduel des taux de sortie sans diplôme entre les enfants de familles immigrées et ceux de familles non immigrées s'élève à 3,2 points. Le niveau en 6ème explique ainsi plus des trois-quarts¹⁵ de l'écart constaté. La différence de niveau au début du collège joue donc bien un rôle primordial dans l'explication des sorties sans diplôme.

Pour le confirmer, on s'appuie sur un indicateur qui mesure l'information qu'apporte le modèle pour « expliquer » le processus de sortie sans diplôme (voir encadré 3). L'indicateur retenu est le critère de Schwartz. Pour le modèle complet (i.e. avec l'ensemble des variables observées), il vaut 10 639. Si on retire les deux scores de la liste des variables retenues par le modèle, le critère passe à 11 591. Si, du modèle complet, on retire toutes les variables sauf les deux scores (c'est-à-dire si on retient uniquement les deux scores), le critère vaut 10 798. Comme l'augmentation du critère consécutive au retrait d'un ensemble de variables

13. $(18,6 - 3,8)/18,6 = 0,796$.

14. $(18,6 - 8,1)/18,6 = 0,564$.

15. $(13,4 - 3,2)/13,4 = 0,761$.

Table 5. Écart brut et résiduels

	Écart brut (%)	Écart résiduel (%)
Sexe de l'élève (<i>ref=garçon</i>)		
<i>filles</i>	-6,8	-4,8***
Structure du ménage (<i>ref=parent isolé</i>)		
<i>parents en couple</i>	-11,3	-5,9***
CS du chef de famille (<i>ref=sans activité</i>)		
<i>agric. exploit., artis., commerc., chefs d'entrep.</i>	-29,6	-5,9***
<i>cadres, prof. intell. sup.</i>	-37,5	-8,6***
<i>prof. intermédiaires</i>	-34,5	-8,5***
<i>employés</i>	-21,9	-3,8*
<i>ouvriers</i>	-18,9	-4,8**
Un parent enseignant? (<i>ref=non</i>)		
<i>oui</i>	-14,3	-4,8***
Diplôme de la mère (<i>ref=non rép., sans objet, aucun ou CEP</i>)		
<i>BEPC, CAP, BEP</i>	-15,3	-5,8***
<i>baccalauréat</i>	-20,8	-6,3***
<i>diplôme ens. supérieur</i>	-24,2	-8,1***
Rapport à la migration (<i>ref=famille non immigrée</i>)		
<i>famille mixte</i>	3,2	4,1***
<i>famille immigrée</i>	13,4	0,5
Le revenu de votre famille est-il ... (<i>ref=très insuffisant ?</i>)		
<i>un peu insuffisant ?</i>	-8,6	-1,9**
<i>juste suffisant ?</i>	-12,0	-2,2***
<i>tout à fait suffisant ?</i>	-17,5	-3,7***
Tranche d'unité urbaine (<i>ref=commune rurale</i>)		
<i>commune urbaine < 5 000 hab</i>	-5,6	-2,5
<i>comm. urb. 5 000-10 000 hab</i>	-6,7	-1,6
<i>comm. urb. 10 000-20 000 hab</i>	-8,6	-3,0
<i>comm. urb. 20 000-50 000 hab</i>	-8,9	-2,5
<i>comm. urb. 50 000-100 000 hab</i>	-7,4	-2,1
<i>comm. urb. 100 000-200 000 hab</i>	-8,0	-2,4
<i>comm. urb. > 200 000 hab</i>	-7,2	-1,6

Seuils de significativité : *** = 1% ; ** = 5% ; * = 10%.

Lecture : Le taux de sortie sans diplôme des enfants vivant avec leurs deux parents est 11,3 points inférieur à celui des enfants de famille monoparentale ; si les enfants des deux types de famille avaient les mêmes caractéristiques (mêmes niveaux en 6ème, ...) et étaient élevés dans des familles ayant les mêmes caractéristiques observées (même catégorie sociale, même diplôme de la mère, ...), l'écart (résiduel) serait de 5,9 points.

Source : DEPP – Panel 1995.

mesure la perte d'information induite, cette perte est donc plus importante avec le retrait des deux scores qu'avec celui de toutes les autres variables. Le niveau en 6ème joue bien le premier rôle.

Pour en quantifier l'importance, on recourt à un exercice de simulation. L'idée est d'estimer la proportion prédite par le modèle de jeunes sortis sans diplôme qui auraient été diplômés s'ils avaient eu les mêmes notes aux tests d'évaluation de 6ème que les élèves qui ont terminé leur scolarité secondaire avec succès, leurs autres caractéristiques restant inchangées. Cet exercice passe par plusieurs étapes, détaillées dans l'encadré 4.

Si leurs scores en 6ème étaient remontés pour être équivalents à ceux des élèves sortis diplômés, 46 % des jeunes constatés sans diplôme seraient en situation d'en avoir un. Autrement dit, et de manière très résumée, le faible niveau au début du collège explique presque la moitié des sorties sans diplôme.

Encadré 4

Niveau en 6ème et sortie sans diplôme

L'objet est de simuler, pour les jeunes repérés comme sortis sans diplôme, des résultats aux tests de 6ème qui les mettent au niveau des jeunes sortis diplômés de l'enseignement secondaire et de calculer, toujours chez ces jeunes non diplômés, la probabilité prédite par le modèle de sortir diplômé, avec cette nouvelle distribution des scores. L'exercice se déroule en trois étapes successives.

Étape 1 : estimation de la propension à sortir sans diplôme

Parmi les élèves terminant leur scolarité secondaire, on sait uniquement distinguer ceux qui sont sortis sans diplôme (autre que le brevet des collèges) et les autres. Or il y a très certainement, chez les jeunes non diplômés, une grande hétérogénéité de situations. Certains ont pu échouer de justesse aux épreuves, d'autres n'avaient pas du tout le niveau. De même, les diplômés ont réussi plus ou moins facilement. On peut ainsi caractériser théoriquement chaque jeune d'une « propension à sortir sans diplôme », plus ou moins élevée, dépendant en particulier des caractéristiques observées. On considèrera, par exemple, qu'un enfant ayant eu de mauvais résultats aux tests de 6ème, dont la mère n'est pas diplômée, etc., aura (probablement) une propension élevée à sortir sans diplôme.

Soit y^* cette propension (inobservée), et x l'ensemble des caractéristiques observées. On suppose que y^* dépend linéairement de x :

$$y^* = x\beta + u$$

où u regroupe toutes les caractéristiques inobservées influant sur y^* . On n'observe pas la propension. On sait juste si le jeune est sorti sans diplôme ($sd = 1$) ou sorti diplômé ($sd = 0$). Entre la propension y^* et la situation du jeune à la sortie du secondaire mesurée par sd , on pose la relation suivante : $sd = 1$ si $y^* > 0$, $sd = 0$ si $y^* < 0$. La sortie sans (resp. avec) diplôme est la manifestation d'une propension supérieure (resp. inférieure) à un certain seuil, fixé sans conséquence dommageable à 0. Dans ces conditions, le calcul pour chaque élève de sa propension à sortir sans diplôme s'effectue de la manière suivante. On suppose que la variable u est distribuée selon la loi normale centrée réduite. On estime les paramètres β et on calcule pour chaque élève i par propension « attendue » $x_i\hat{\beta}$. On lui impute une propension individuelle

$x_i\hat{\beta} + u_i$ en tirant une valeur de u dans la loi normale centrée réduite qui vérifie $x_i\hat{\beta} + u_i > 0$ si l'élève n'est pas diplômé ($sd = 1$) ou $x_i\hat{\beta} + u_i < 0$ s'il est diplômé ($sd = 0$).

Étape 2 : simulation des scores aux tests de 6ème pour les non diplômés

L'objectif est de simuler une distribution des scores chez les non diplômés qui soit la plus proche possible de celle constatée chez les diplômés. On fait en sorte que les deux distributions aient la même moyenne et le même écart-type. L'exercice est le même pour les scores en mathématiques et pour ceux en français.

Soit \bar{s}_{sd} et \bar{s}_d les moyennes des scores constatées respectivement chez les sans diplôme et les diplômés. Soit σ_{sd} et σ_d les écarts-type correspondants. Soit s_i le score constaté de l'élève i sorti sans diplôme. On lui affecte le score \tilde{s}_i :

$$\tilde{s}_i = \frac{\sigma_d}{\sigma_{sd}}(s_i - \bar{s}_{sd}) + \bar{s}_d$$

On vérifie que la moyenne et l'écart-type des \tilde{s}_i , calculés sur la sous-population des sans diplôme, valent \bar{s}_d et σ_d .

Étape 3 : proportion des sans diplôme « devenant » diplômés

Connaissant pour chaque jeune sans diplôme ses « nouveaux » scores, on recalcule sa propension à sortir sans diplôme $x_i\hat{\beta} + u_i$ en remplaçant les valeurs constatées du score en mathématiques et du score en français par celles simulées en étape 2. Puis on calcule la proportion de ceux qui vérifient $\tilde{x}_i\hat{\beta} + u_i < 0$, où $\tilde{x}_i = x_i$ sauf pour les deux variables de scores.

5. Poursuivre l'analyse

L'objectif de l'étude était de mettre en évidence les caractéristiques sociodémographiques distinguant les sortants sans diplôme des autres et, ce faisant, d'identifier en première analyse les facteurs explicatifs du processus de sortie sans diplôme. Son résultat principal est le rôle prédominant du niveau en 6ème. Cela rejoint un constat largement partagé sur le fait que le décrochage scolaire est l'aboutissement d'un processus qui trouve très souvent son origine dans les premières années de scolarisation¹⁶. D'où la pertinence et la nécessité d'une politique de prévention du décrochage menée très en amont dans le parcours scolaire.

Si le niveau à l'entrée au collège détermine assez fortement le décrochage, qu'est-ce qui détermine alors ce niveau? L'attention est donc déportée sur l'enseignement primaire, et le rôle des caractéristiques sociodémographiques doit être réexaminé. Notamment, l'étude a conclu que la catégorie sociale des parents avait somme toute un rôle relativement limité par rapport à la variable de niveau. Mais le niveau en 6ème est lui-même (partiellement) déterminé par le milieu social de l'élève et son environnement familial. Rien ne dit alors que les inégalités sociales ou toute autre forme d'inégalités ne jouent pas, *tous comptes faits*, un rôle déterminant en la matière.

La feuille de route de travaux complémentaires est en partie écrite, consistant à suivre des cohortes d'élèves entrés en cours préparatoire une année donnée, grâce aux dispositifs

16. Voir à ce titre : Éric Maurin (2012), *Quelles politiques contre le décrochage et le déclassé ?* Formation Emploi, CEREQ, n° 118.

d'observation mis en place par la DEPP et analogues à la source utilisée pour les besoins de la présente étude. L'analyse des parcours scolaires de ces enfants, celle des facteurs contribuant au creusement des inégalités à l'école ou des facteurs favorisant au contraire leur aplanissement, fait certainement partie des priorités à inscrire sur l'agenda d'études ultérieures.

Annexe 1 – Les effectifs de l'échantillon

	Sortants sans diplôme	Sortants avec diplôme
	2 162	11 301
Sexe de l'élève		
<i>filles</i>	871	5 858
<i>garçons</i>	1 291	4 443
Structure du ménage		
<i>parent isolé</i>	440	1 337
<i>parents en couple</i>	1 722	9 964
CS du chef de famille		
<i>sans activité</i>	94	135
<i>agric. exploit., artis., commerc., chefs d'entrep.</i>	215	1 449
<i>cadres, prof. intell. sup.</i>	122	1 985
<i>prof. intermédiaires</i>	209	2 268
<i>employés</i>	444	1 818
<i>ouvriers</i>	1 078	3 646
Un parent enseignant ?		
<i>oui</i>	47	1 064
<i>non</i>	2 115	10 237
Diplôme de la mère		
<i>non réponse, sans objet, aucun ou CEP</i>	1 222	3 182
<i>BEPC, CAP, BEP</i>	637	4 049
<i>baccalauréat</i>	163	1 711
<i>diplôme ens. supérieur</i>	140	2 359
Rapport à la migration		
<i>famille non immigrée</i>	1 676	9 761
<i>famille mixte</i>	126	584
<i>famille immigrée</i>	360	956
Pensez-vous que le revenu de votre famille est		
<i>très insuffisant ?</i>	675	2 028
<i>un peu insuffisant ?</i>	597	2 925
<i>juste suffisant ?</i>	688	4 263
<i>tout à fait suffisant ?</i>	202	2 085
Tranche d'unité urbaine de la commune de l'étab.		
<i>commune rurale</i>	90	295
<i>commune urbaine < 5 000 hab</i>	78	361
<i>comm. urb. 5 000-10 000 hab</i>	147	734
<i>comm. urb. 10 000-20 000 hab</i>	175	1 000
<i>comm. urb. 20 000-50 000 hab</i>	288	1 685
<i>comm. urb. 50 000-100 000 hab</i>	265	1 385
<i>comm. urb. 100 000-200 000 hab</i>	233	1 267
<i>comm. urb. > 200 000 hab</i>	886	4 574

Lecture : 871 observations correspondent à des filles qui sont sorties sans diplôme.

Source : DEPP – Panel 1995.