

Ministère de l'éducation nationale, de la jeunesse et des sports

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

Document de travail

Série « Méthodes »

N°2021-M02

InserJeunes – taux d'interruption en cours de formation

Loïc Midy

Ce document décrit la méthodologie du taux d'interruption en cours de formation.

Le taux d'interruption en cours de formation est un des indicateurs calculés par InserJeunes pour répondre à la loi du 5 septembre 2018 pour la liberté de choisir son avenir professionnel. Les établissements font ici référence aux centres de formation d'apprentis (CFA) et aux lycées professionnels. De même, le terme élèves est général et englobe également les apprentis. Les apprentis qui font leur formation dans un lycée professionnel sont comptabilisés d'une part pour établir les statistiques de leur CFA et d'autre part pour établir les statistiques du lycée professionnel dans lequel ils étudient.

Cet indicateur donne une estimation du risque d'interrompre sa formation sur l'ensemble de la durée du diplôme. L'information utilisée pour estimer une interruption de formation une année de scolarité donnée est toujours la plus récente disponible.

L'indicateur n'est pas disponible au niveau France. Il est calculé uniquement par diplôme et par établissement et seulement pour les diplômes suivants CAP, Mention complémentaires de niveau V, Mentions complémentaires de niveau IV, Brevet professionnel (BP) pour les apprentis, bac pro et BTS, pour lesquels nous disposons des bases de réussite aux examens des élèves. Comme pour les autres indicateurs InserJeunes au niveau établissement on cumule les données de deux années scolaires. Ces statistiques sont diffusés sur le site du ministère de l'éducation nationale en open data¹.

Le taux d'interruption en cours de formation est calculé en utilisant une méthode conjoncturelle similaire à celle utilisée pour calculer le taux d'accès au baccalauréat qui est un des indicateurs de valeur ajoutée (IVAL) établis par la DEPP de longue date.

Sommaire

Contenu

1. Présentation synthétique de l'indicateur taux d'interruption en cours de formation.....	3
1.1. Définition.....	3
1.2. Présentation synthétique du calcul du taux d'interruption en cours de formation.....	3
2. Détail de la méthode de calcul	4
3. Justification de la méthode de calcul retenue	6
3.1. Calculer un taux d'interruption en cours de formation par une méthode par suivi de cohorte ou une méthode conjoncturelle ?.....	6
3.2. Est-ce que les deux méthodes donnent des résultats identiques ?.....	6
4. Résultats InserJeunes sur les sortants de l'été 2018 et 2019 cumulés	8

¹ Ici <https://www.education.gouv.fr/l-insertion-des-jeunes-apres-une-formation-en-voie-professionnelle-307956>

1. Présentation synthétique de l'indicateur taux d'interruption en cours de formation

1.1. Définition

L'indicateur mesure le risque qu'un élève qui entre dans une formation donnée interrompe sa formation avant d'être diplômé. Un élève est considéré comme interrompant sa formation si :

- étant en dernière année de formation, il n'obtient pas son examen et ne redouble pas l'année suivante.
- étant en cours de formation, il ne passe pas dans l'année supérieure dans le même groupe de spécialité de formation du même diplôme, ne redouble pas et n'a pas obtenu son examen de façon anticipée.

1.2. Présentation synthétique du calcul du taux d'interruption en cours de formation

Prenons l'exemple d'une formation en 2 ans, par exemple des CAP en 2 ans.

Comme pour les autres indicateurs InserJeunes on cumule les données de deux années scolaires.

Exemple illustratif : supposons que l'on ait la répartition des élèves en CAP suivante dans un établissement

	année formation	inscrits	succès	doublants	taux passage
inscrits en 2017-2018	1	110	95	0	
	2	115	90	3	
inscrits en 2018-2019	1	100	85	5	
	2	100	85	2	
cumul 2 ans	1	210	180	5	87,8%
	2	215	175	5	83,3%

Etape 1 : On compte le nombre **d'élèves** par année de formation pour chacune des 2 années scolaires puis on cumule les données. Dans cet exemple, on a 210 inscrits en 1^{re} année de formation en cumulant le nombre d'inscrits en 1^{re} année de formation en 2017-2018 **plus** le nombre d'inscrits en 1^{re} année de formation en 2018-2019.

Etape 2 : Pour chaque élève, on regarde l'année scolaire suivante sa situation. Il peut alors :

- être redoublant
- être « **en succès** », **c'est-à-dire** passer dans l'année suivante de formation dans **le même groupe de spécialité de formation** ou **réussir l'examen final de la formation**.
- avoir **interrompu** cette formation

Par exemple, un apprenti en 1^{re} année de CAP dans une spécialité communication et information qui est l'année suivante en 2^e année de CAP dans une spécialité services aux personnes est comptabilisé comme ayant interrompu sa formation initialement suivie (car il change de groupe de spécialité de formation).

Etape 3 : On calcule un taux de passage pour chaque année de formation sur les données cumulées en appliquant la formule $\text{taux de passage} = \frac{\text{succès}}{\text{inscrits} - \text{doublants}}$.

Dans l'exemple présenté ci-dessus,

- le taux de passage 1->2 = $180 / (210-5) = 87,8 \%$
- le taux de passage 2->examen passé = $175 / (215-5) = 83,3 \%$

Etape 4 : Le taux d'interruption en cours de formation est alors égal à :

1 – taux d'accès global soit :

1 – (taux passage 1->2) X (taux passage 2-> examen obtenu)

Dans l'exemple présenté ci-dessus, le taux d'abandon en cours de formation vaut :

1 – 87,8% X 83,3%

Soit un taux d'interruption en cours de formation de 27 %

2. Détail de la méthode de calcul

Etape 1 : calcul du devenir de chaque élève

On calcule le devenir de chaque élève/apprentis du champ InserJeunes l'année scolaire suivante en notant les élèves en succès et les élèves redoublants.

Ce suivi des élèves selon la formation est réalisé de manière globale sur les apprentis et la voie professionnelle scolaire MENJ et agricole. Par exemple, un élève de seconde pro en voie pro sco qui poursuit sa formation en première en apprentissage est comptabilisé comme un succès.

On ne recherche le devenir des élèves/apprentis l'année scolaire suivante que dans une formation en apprentissage ou en voie professionnelle scolaire MENJ (dans le public ou le privé sous contrat) ou agricole.

Un élève **inscrit** dans une formation pour une année de formation donnée n-1/n (ex : 1^{re} année de bac pro soit une seconde pro) est en **succès** s'il est dans l'un des 2 cas suivants :

- **cas 1** : il a la même grande catégorie de formation en n-1/n et n/n+1 (ces catégories sont CAP, BTS, MC5, MC4, BP, bac pro) **ET** il a même 2 premières positions de groupe NSF (nomenclature N_GROUPE_SPECIALITE de la BCN²) en n-1/n et n/n+1 **ET** Année de formation n/n+1 = (Année de formation n-1/n) **+ 1**
- **cas 2** : il a réussi l'examen à la fin de l'année scolaire n-1/n

Nb : Un élève qui est dans le cas 1 ou le cas 2 mais qui change d'établissement est bien noté en succès.

Dans le dispositif IVAL, qui pour le moment ne s'applique que pour le baccalauréat voie scolaire, un élève inscrit est en succès s'il passe dans la classe supérieure dans le même établissement ou s'il réussit son examen.³

Un élève **inscrit** dans une formation pour une année de formation donnée n-1/n (ex : 1^{re} année de bac pro soit une seconde pro) est **redoublant** s'il a la même grande catégorie de formation en n-1/n et n/n+1 (ces catégories sont CAP, BTS, MC5, MC4, BP, bac pro) **ET** il a même 2 premières positions de groupe NSF (nomenclature N_GROUPE_SPECIALITE de la BCN) en n-1/n et n/n+1 **ET** Année de formation n/n+1 = Année de formation n-1/n

Etape 2 : calcul des taux de passage élémentaires

On calcule des taux de passage élémentaires selon la maille de calcul établissement * niveau de diplôme * durée du cursus * année de formation. Les données sont cumulées selon les 2 dernières années scolaires disponibles.

Exemple illustratif (toujours le même exemple) : supposons que l'on ait la répartition des élèves en CAP suivante dans un établissement

	année formation	inscrits	succès	doublants
inscrits en 2017-2018	1	110	95	0
inscrits en 2018-2019	1	100	85	5

² Cette nomenclature est diffusée notamment ici

http://infocentre.plejade.education.fr/bcn/workspace/viewTable/n/N_GROUPE_SPECIALITE

³ page 101 de l'article disponible en téléchargement ici <https://www.education.gouv.fr/media/14849/download>

cumul 2 ans	1	210	180	5
--------------------	----------	------------	------------	----------

Alors j'ai au total sur le cumul des 2 années, $110+100=210$ inscrits, $95+85=180$ succès et $0+5=5$ redoublants et un taux de passage de 1^{re} à 2^e année de $180/(210-5)$ soit 87,8%.

Etape 3 : suppression des données lorsque les effectifs sont insuffisants

Pour toutes les mailles de calculs pour lesquels l'effectif d'inscrits est inférieur à 20 on supprime toutes les lignes qui ont même (établissement, niveau de diplôme, durée du cursus).

Exemple illustratif) : supposons que l'on ait la répartition des élèves en bac pro suivante dans un établissement

	année formation	inscrits
cumul 2 ans	1	50
cumul 2 ans	2	50
cumul 2 ans	3	16

Alors dans ce cas de figure on ne diffusera pas de taux d'interruption en cours de formation pour le bac pro de cet établissement (croisement : cet établissement X bac pro).

Lorsqu'on calcule le taux d'interruption en cours de formation pour 2017-2018 et 2018-2019 cumulés cela conduit à supprimer 4 153 lignes sur un total de 22 746 soit 18% des lignes⁴. On conserve donc alors $22\ 746 - 4\ 153 = 18\ 593$ lignes.

Etape 4 : suppression des données lorsque le taux d'appariement avec les bases examens est trop faible

Pour obtenir une estimation fiable du taux d'interruption en cours de formation il faut que la majorité des élèves qui passent l'examen soient appariés avec les bases de résultats aux examens. Etant donné que nous ne disposons pas toujours de l'INE dans les bases de résultats aux examens nous réalisons également des appariements sur les noms, prénoms et date de naissance des élèves par nature moins précis que l'appariement sur l'INE.

Nous avons fixé le seuil de qualité à un taux d'appariement minimal de 75%.

Lorsqu'on calcule le taux d'interruption en cours de formation pour 2017-2018 et 2018-2019 cumulés cela conduit à supprimer 683 lignes sur un total de 18 593 soit 3,7% des lignes.

Etape 5 : correction des données spécifiques pour les centres de formations d'apprentis (CFA)

Si, pour un CFA et une catégorie d'examen donnée, on dispose de taux d'accès pour toutes les années sauf la première année de formation alors on ajoute un taux d'accès 1^{re} à 2^e année avec la valeur 1. On fait cette correction pour prendre en compte le cas des CFA qui préparent des apprentis au bac uniquement en première et terminale (la seconde étant faite hors apprentissage).

Cela conduit à ajouter 64 taux d'accès de 1^{re} à 2^e année à 1. En effet, on observe des parcours de formation mixtes avec notamment un début de formation en voie professionnelle scolaire suivie d'une fin de formation en apprentissage (cf. note d'information Les réorientations dans l'enseignement professionnel sont majoritairement de la voie scolaire vers l'apprentissage⁵).

⁴ L'année suivante le chiffre devrait être plus bas car pour ce millésime uniquement on ne disposait des données que d'un millésime pour la voie pro sco agricole. Il n'y a pas eu de cumul sur 2 ans donc les effectifs d'inscrits étaient plus faibles pour ce champ et par conséquent le nombre de cases éliminées était plus élevé qu'en régime de croisière normal.

⁵ <https://www.education.gouv.fr/les-reorientations-dans-l-enseignement-professionnel-sont-majoritairement-de-la-voie-scolaire-vers-l-309358>

Si, pour un CFA et une catégorie d'examen donnée, on dispose de taux d'accès pour toutes les années sauf la dernière année de formation et que pour ce croisement on n'a pas déjà fait la correction ¹⁶ alors on ajoute un taux d'accès avant-dernière à dernière année avec la valeur 1. Cela conduit à ajouter 7 taux d'accès d'avant-dernière à dernière année à 1.

Etape 6 : suppression des données lorsqu'il manque un taux de passage élémentaire

Si pour un établissement, catégorie d'examen il manque au moins un taux d'accès élémentaire alors il n'est pas possible de calculer de taux d'accès global, ce dernier étant le produit des taux d'accès élémentaires.

Cela conduit à ne pas calculer de taux d'accès global dans 22 sur 7867 soit 0,3% des cas.

Etape 7 : On calcule les taux d'interruption en cours de formation

On applique la formule taux d'interruption en cours de formation = 1 – taux d'accès global

Etape 8 : suppression des taux d'interruption en cours de formation très élevés

On supprime les taux d'interruption supérieur à 90% ce qui conduit à supprimer 12 cas 7748 soit 0,2% des cas. On considère qu'il s'agit d'anomalies.

3. Justification de la méthode de calcul retenue

3.1. Calculer un taux d'interruption en cours de formation par une méthode par suivi de cohorte ou une méthode conjoncturelle ?

Pour calculer le taux d'interruption en cours de formation il est possible d'utiliser soit une méthode par suivi de cohorte soit une méthode conjoncturelle.

Un suivi de cohorte consiste à suivre un ensemble d'individus qui ont vécu un événement semblable au cours d'une période de temps (habituellement une année). Dans notre cas, elle consiste à faire le suivi de l'ensemble des élèves dès leur entrée dans un cursus donné (ils seront appelés primo inscrits par la suite). Si l'on reprend le cas des CAP en 2 ans, cela veut dire qu'on part du champ des élèves entrant en première année de CAP une année scolaire donnée (ce champ ne comprend donc pas les redoublants) et qu'on regarde leur devenir une fois que tous les élèves ont achevé ou interrompu leur formation. Par exemple s'il y a 100 inscrits en première année de CAP en 2017-2018 et si 3 ans plus tard 80 d'entre eux ont obtenu leur CAP alors le taux d'interruption est de $(100-80)/100 = 20\%$.

L'inconvénient de cette méthode est qu'il faut attendre que tous les élèves aient achevé ou interrompu leur formation y compris les redoublants. Donc pour un bac pro donné les résultats d'une année scolaire donnée sont connus 4 ans plus tard (on compte les 3 ans du cursus et on ajoute une année supplémentaire ce qui permet de tenir compte de la majorité⁷ de ceux qui ont redoublé à un moment du cursus).

Nous avons considéré que ces délais n'étaient pas satisfaisants ce qui nous a conduit à retenir pour InserJeunes la méthode conjoncturelle présentée précédemment.

3.2. Est-ce que les deux méthodes donnent des résultats identiques ?

Si le phénomène étudié est « stationnaire » c'est-à-dire que tous les flux d'élèves sont constants toutes les années (dans l'exemple simplifié infra d'une formation en 1 an cela veut dire notamment qu'il y a chaque année 100 primo inscrits dont 80 seront en succès au bout d'un 1 an, 9 redoubleront

⁶ car on ne veut pas imputer plus d'un taux d'accès pour un croisement CFA, catégorie d'examen.

⁷ Mais pas tous : si un élève redouble 2 fois pendant le cursus de bac pro alors au bout de 4 ans il n'a pas encore achevé sa formation.

et 11 abandonneront) pendant une échelle de temps correspond au parcours de la cohorte (soit 4 ans maximum pour la formation bac pro), alors les deux méthodes donnent exactement le même résultat.

Nous illustrons ce résultat à travers un exemple simplifié d'une formation en 1 an. Pour les besoins de pédagogie, des nombres non entiers ont été introduits.

année scolaire	Primo inscrits	redoublants	ensemble
2017-2018	100	9	109
devenir des élèves en 2018-2019			
Succès	80	8	88
Redoublants	9	0	9
Abandons	11	1	12

Méthode suivi de cohorte

On considère une cohorte de primo entrants une année donnée.

Ainsi, parmi les 100 primo entrants dans la formation l'année scolaire 2017-2018 on a :

- 80 succès sans redoublement.
- 9 redoublants et parmi eux on a encore 8 succès.

Ainsi le taux d'accès est $(80 + 8) / 100 = 88\%$ et le taux d'abandon est 12%.

Méthode conjoncturelle

Dans cette méthode, on considère l'ensemble des élèves d'une année scolaire donnée (primo entrants et redoublants)

Taux d'accès = $\frac{\text{succès}}{\text{inscrits} - \text{doublants}}$ est de $88 / (109 - 9) = 88\%$ et le taux d'abandon est 12%. Les deux méthodes donnent donc bien le même résultat.

Décomposons le taux d'accès conjoncturel :

- Le numérateur doit refléter tous les succès des élèves qu'ils aient ou non redoublé soit le 88 de la colonne ensemble.
- Le dénominateur doit refléter uniquement les primo- entrants donc on fait inscrits moins doublants soit $109 - 9 = 100$ et on retombe bien sur l'effectif de primo inscrits.

3.3. Justification du niveau de finesse de formation retenue pour statuer sur un succès ou non

Dans un premier temps nous avons réalisé le suivi des élèves en retenant la formation fine sur 5 positions soit le groupe NSF sur 3 positions (nomenclature N_GROUPE_SPECIALITE de la BCN) plus 2 positions supplémentaires mais cela aurait conduit à ne pas considérer certains élèves en succès alors que visiblement ils poursuivent bien la même formation par exemple dans les cas suivants :

Libellé de la formation suivie en classe de seconde	Libellé de la formation suivie l'année suivante en classe de première
2NDPRO ACC.SOINS-SERV.PERS. 2NDE COMMUNE	1ERPRO ACC.SOINS-S.PERS. OPT.EN STRUCTUR
2NDPRO MAINTENANCE VEHIC. 2NDE COMMUNE	1ERPRO MAINT.VEHIC.OPTA VOIT.PARTICUL.
2NDPRO SYSTEMES NUMERIQUES 2NDE COMMUNE	1ERPRO SYST.NUM.OPT.C RES.INF.SYST.COMM.
2NDPRO ACC.SOINS-SERV.PERS. 2NDE COMMUNE	1ERPRO ACC.SOINS-S.PERS. OPT.A DOMICILE

Dans un second temps nous avons réalisé le suivi des élèves en retenant le groupe NSF sur 3 positions (nomenclature N_GROUPE_SPECIALITE de la BCN mais cela aurait encore conduit à ne

pas considérer certains élèves en succès alors que visiblement ils poursuivent bien la même formation spécifiquement pour la voie professionnelle agricole.

Au final, nous avons donc retenu les 2 premières positions de groupe NSF (nomenclature N_GROUPE_SPECIALITE de la BCN) ce qui ne change quasiment pas les statistiques pour l'apprentissage et la voie professionnelle champ éducation nationale par rapport au suivi sur les 3 positions de groupe NSF et permet de bien détecter tous les succès pour la voie professionnelle agricole.

4. Résultats InserJeunes sur les sortants de l'été 2018 et 2019 cumulés

Voici les principaux quantiles du taux d'interruption en cours de formation pour les années 2017-2018 et 2018-2019 cumulées

	CFA			lycée professionnels		
	bac pro	BTS	CAP	bac pro	BTS	CAP
100Max	87,4%	89,7%	88,6%	84,0%	85,5%	85,5%
100%						
99%	76,7%	76,4%	73,8%	63,7%	64,7%	63,0%
95%	65,1%	61,0%	54,0%	56,4%	53,1%	54,6%
90%	57,4%	55,5%	47,5%	51,3%	47,6%	49,6%
75% Q3	49,4%	45,0%	41,3%	44,2%	40,3%	42,8%
50%						
Médiane⁸	41,2%	35,2%	34,5%	36,0%	32,5%	34,6%
25% Q1	32,2%	27,4%	27,5%	27,9%	24,8%	26,0%
10%	24,1%	19,6%	20,0%	21,1%	18,3%	18,7%
5%	19,8%	13,7%	15,6%	17,5%	14,6%	14,6%
1%	10,2%	8,1%	5,8%	10,3%	8,1%	8,4%
0% Min	4,3%	4,3%	0,0%	3,6%	0,0%	0,0%

Note de lecture : pour la formation bac pro, le taux d'interruption en cours de formation est supérieur ou égal à 65,1% pour 5% des CFA. Pour la formation bac pro, le taux d'interruption en cours de formation est supérieur ou égal à 56,4 % pour 5% des lycées professionnels.

On rappelle que l'on a supprimé les taux d'interruption supérieur à 90% (soit 12 cas). Toutefois, on observe qu'il y a peu de établissements qui ont des taux d'interruption pour les niveaux bac pro, BTS ou CAP très élevés. En effet, ces taux sont supérieurs à 75% dans seulement 1% des cas pour les CFA et supérieur à 65% dans seulement 1% des cas pour les lycées professionnels.

⁸ Si on ordonne une distribution, la médiane partage cette distribution en deux parties d'effectifs égaux.