

**Archive ouverte relative à l'enquête : « un premier bilan de l'accès à l'enseignement supérieur dans le cadre de la loi orientation et réussite des étudiants »**

**Éléments associés à l'annexe 4 (« Sondages d'opinion mis en miroir »)**

Pour chacun des deux catégories de sondages (triple sondage candidats et sondage auprès des professeurs principaux en classe de Terminale), les réponses sont proposées sur l'archive en ligne sous deux formats :

- une synthèse transversale sous la forme d'un rapport de synthèse détaillant les principaux résultats mobilisés dans le rapport ;
- l'ensemble des résultats bruts anonymes correspondant à l'assiette retenue pour permettre au citoyen intéressé de réaliser ses propres analyses : visualisations des données, tris croisés, analyses corrélatives, régressions, partitionnement de données, etc.

**Éléments associés à l'annexe 5 (« Des réorientations toujours plus nombreuses et onéreuses »)**

L'ensemble des hypothèses et tableaux intermédiaires permettant de reproduire les résultats présentés dans le tableau de synthèse en annexe 5 du rapport sont proposés sur l'archive en ligne.

Pour le détail du calcul des coûts de formation pour l'État et les étudiants : voir les tableaux et l'encadré *infra*.

**Tableau n° 1 : Estimation du coût unitaire moyen de formation pour l'État en 2016 en STS, CPGE et pour l'ensemble des établissements de l'enseignement supérieur.**

	STS	CPGE	Ensemble des établissements du second degré	Universités	Ensemble des établissements du supérieur	Hypothèses et clés de lecture
<i>Coût annuel total par étudiant [C]</i>	13 780	15 110	9 720	10 210	11 510	
<i>Part des dépenses consacrées à l'enseignement supérieur et à la recherche (ESR) [P1]</i>	87,7 %					Hypothèse : la part des dépenses consacrées à l'ESR est supposée identique pour l'ensemble des formations sous revue.
<i>Part des dépenses d'ESR consacrées aux activités d'enseignement [P2]</i>	100 %	100 %	100 %	50 %	50 % à 100 %	Le taux de 50 % correspond à la moitié d'un service d'un enseignant-

						chercheur à l'université.
<i>Part du financement assuré par l'État [P3]</i>	64,3 % à 73,2 %	64,3 % à 73,2 %	64,3 % à 73,2 %	81,6 %	9,5 % à 75,5 %	<p>Clé de lecture : les fourchettes basses et hautes correspondent respectivement au privé subventionné et au public.</p> <p>Hypothèse pour les STS et CPGE : en l'absence de données détaillées, la part de financement assurée par l'État est supposée la même que pour l'ensemble des établissements du second degré.</p>
<i>Coût unitaire moyen de formation pour l'État (en euros) [= C × P1 × P2 × P3]</i>	Privé subventionné : 7 771  Public : 8 846  Moyenne : 8 309	Privé subventionné : 8 521  Public : 9 700  Moyenne : 9 111	Privé subventionné : 5 481  Public : 6 240  Moyenne : 5 861	3 653	Privé subventionné : 479 à 959 (Moyenne : 719)  Public : 3 809 à 7 618 (Moyenne : 5 714)	<p>Clé de lecture : pour les universités, le chiffrage obtenu avec cette méthode de calcul est cohérent avec le montant estimé par la Mission expertise et conseil de la DGESIP : de 3 645 euros pour la PACES à 4 800 euros pour les licences hors-PACES (voir <b>Erreur ! Source du renvoi introuvable..</b></p>

Source : retraitements de la Cour des comptes à partir des « Repères et références statistiques » (2018) de la DEPP du MENJ-MESRI, page 325 (item 1 pour C, item 2 pour P1 et item 3 pour P2) et page 323 (items 1 et 2 pour P3).

**Tableau n° 2 : Évolution des frais de dossiers moyens (en euros) acquittés par les étudiants suivant leur filière d'étude.**

<i>Type de formation</i>	<b>Parcoursup 2018</b>		<b>Parcoursup 2019</b>	
	Non-boursiers	Boursiers	Non-boursiers	Boursiers
<i>Autres formations</i>	46	29	68	15
<i>BTS - BTSA</i>	15	15	15	15
<i>CPGE</i>	15	15	16	16
<i>Cycle universitaire préparatoire aux grandes écoles</i>	NE	NE	40	40
<i>DCG</i>	15	15	16	16
<i>DEUST</i>	NE	NE	130	130
<i>Écoles d'architecture</i>	39	4	39	2
<i>Écoles de commerce (et de management)*</i>	77	52	100	39
<i>Écoles supérieures d'art</i>	45	41	47	43
<i>Formations d'ingénieurs</i>	83	2	91	3
<i>Formations du social</i>	138	138	NE	NE
<i>Formations paramédicales</i>	60	60	NE	NE
<i>Formations diplômantes du secteur sanitaire et social (dont IFSI et EFTS)</i>	NE	NE	123	121
<i>Formations préparatoires à l'enseignement supérieur</i>	NE	NE	40	40
<i>Licence(s)*</i>	102	102	81	72
<i>Mentions complémentaires</i>	15	15	15	15

*Source : Cour des comptes à partir des données de Parcoursup 2018 et 2019 (juin). Pour un type de formation donnée (variable g\_tf\_lib transformée pour inclure la PACES), les frais moyens ont été calculés en tenant compte du poids relatif de chaque formation concernée en termes de capacités d'accueil (moyenne des frais pondérée par les capacités). (\*) Terminologie entre parenthèse ajoutée ou supprimée dans Parcoursup 2019. NE : catégorie inexistante pour l'année concernée.*

**Tableau n° 3 : Évolution des frais de dossiers moyens (en euros) acquittés par les étudiants suivant le type de contrat de leur établissement.**

<i>Type de contrat avec l'État</i>	<b>Parcoursup 2018</b>		<b>Parcoursup 2019</b>	
	Non-boursiers	Boursiers	Non-boursiers	Boursiers
<i>Formations publiques du secondaire et du supérieur</i>	66	4	68	3
<i>Formations privées sous contrat du secondaire</i>	94	3	111	54
<i>Formations privées hors contrat du secondaire</i>	44	36	146	145
<i>Formations privées du supérieur</i>	79	31	91	37

*Source : Cour des comptes à partir des données d'APB et Parcoursup. Pour un type de contrat d'établissement donné (variable « g\_tc\_lib »), les frais moyens ont été calculés en tenant compte du poids relatif de chaque établissement concerné en termes de capacités d'accueil (moyenne des frais pondérée par les capacités).*

**Tableau n° 4 : Évolution des frais d'inscriptions moyens annuels (en euros) acquittés par les étudiants suivant leur filière d'étude.**

<i>Type de formation</i>	<b>Parcoursup 2018</b>		<b>Parcoursup 2019</b>	
	Non-boursiers	Boursiers	Non-boursiers	Boursiers
<i>Autres formations</i>	3 194	2 937	2 625	1982
<i>BTS - BTSA</i>	1 318	1 243	1 314	1235
<i>CPGE</i>	2 110	2 098	2 159	2114
<i>Cycle universitaire préparatoire aux grandes écoles</i>	186	0	172	0
<i>DCG</i>	1 204	956	1 154	958
<i>DEUST</i>	184	0	167	0
<i>Diplômes des métiers d'Arts</i>	1 741	1 741	NE	NE
<i>DN MADE</i>	1 574	1 574	2 210	2 511
<i>DU</i>	1 807	557	894	521
<i>DUT</i>	185	1	171	1
<i>Écoles d'architecture</i>	471	83	469	92
<i>Écoles de commerce (et de management)*</i>	6 285	6 212	6 311	6 994
<i>Écoles supérieures d'art</i>	589	384	584	358
<i>Formation en ingénierie</i>	184	0	170	0
<i>Formations d'ingénieurs</i>	3 887	3 913	4 840	4 240
<i>Formations du social</i>	1 612	925	NE	NE
<i>Formations paramédicales</i>	1 088	745	NE	NE

<i>Type de formation</i>	<b>Parcoursup 2018</b>		<b>Parcoursup 2019</b>	
<i>Formations diplômantes du secteur sanitaire et social (dont IFSI et EFTS)</i>	NE	NE	606	338
<i>Licences</i>	404	138	408	155
<i>Mentions complémentaires</i>	783	760	974	892
<i>Métiers du sport</i>	NE	NE	188	185
<i>PACES</i>	184	0	170	0

Source : Cour des comptes à partir des données de Parcoursup 2018 et 2019 (juin). Pour un type de formation donné, les frais moyens ont été calculés en tenant compte du poids relatif de chaque formation concernée en termes de capacités d'accueil (moyenne des frais pondérée par les capacités). Les frais d'inscription, renseignés dans un format textuel par les responsables de chaque formation, ont pu être reconstruits dans un format exploitable numériquement au travers d'une analyse sémantique (voir encadré infra). (\*) Terminologie entre parenthèse ajoutée dans Parcoursup 2019. NE : catégorie inexistante pour l'année concernée.

**Tableau n° 5 : Évolution des frais d'inscriptions moyens annuels (en euros) acquittés par les étudiants suivant leur filière d'étude.**

<i>Type de formation</i>	<b>Parcoursup 2018</b>		<b>Parcoursup 2019</b>	
	Non-boursiers	Boursiers	Non-boursiers	Boursiers
<i>Formations publiques du secondaire et du supérieur</i>	225	31	229	48
<i>Formations privées sous contrat du secondaire</i>	2 011	1 929	2 004	1 874
<i>Formations privées hors contrat du secondaire</i>	3 568	3 568	3 235	3 043
<i>Formations privées du supérieur</i>	5 816	5 059	5 881	5 269

Source : Cour des comptes à partir des données de Parcoursup 2018 et 2019 (juin). Pour un type de contrat d'établissement donné (variable « g\_tc\_lib »), les frais moyens ont été calculés en tenant compte du poids relatif de chaque établissement concerné en termes de capacités d'accueil (moyenne des frais pondérée par les capacités). Les frais d'inscription, renseignés dans un format textuel par les responsables de chaque formation, ont pu être reconstruits dans un format exploitable numériquement au travers d'une analyse sémantique (voir encadré infra).

### **L'analyse des frais de scolarité pour les formations d'APB et Parcoursup par le biais d'une méthode sémantique**

Les montants des frais de scolarité ont été renseignés par les formations sur APB et Parcoursup sous forme de textes descriptifs. Pour intégrer ces éléments à l'analyse, un processus de récupération automatique du montant annuel des frais liés à la scolarité a été mis au point. L'algorithme, écrit en langage python, repose sur deux éléments clefs, qui sont : la capacité à repérer un chiffre dans un élément de texte et la capacité à donner du sens aux chiffres récupérés.

Le premier point a été résolu grâce au package « re » de python (pour *regular expressions*) : ce package permet, au prix d'un peu de paramétrage, de repérer des chaînes de caractère et donc notamment d'identifier des nombres décimaux.

Le deuxième point s'est avéré plus problématique. L'absence généralisée de format empêche en effet l'élaboration d'un algorithme de lecture des frais de scolarité vraiment efficace. Quelques-unes de ces difficultés ont pu être dépassées grâce à la constitution de dictionnaires d'interprétation qui guident la lecture des libellés descriptifs des frais de formation.

Certains problèmes d'interprétation n'ont cependant pas pu être résolus, qui sont largement liés à des modes de saisie fautifs et peuvent même parfois gêner un candidat lors de l'étape de la formulation des vœux :

- Certaines formations n'ont pas renseigné la case « frais de scolarité ».
- En ce qui concerne les frais de scolarité des boursiers, certaines formations n'ont utilisé que du texte pour signifier que la scolarité était gratuite : certaines informations ont donc peut-être été perdues par l'algorithme.
- Toutes les formations ne prennent pas en compte les mêmes éléments dans leur calcul du montant des frais de scolarité : certaines formations n'indiquent que les frais liés au suivi des cours, d'autres comptent des postes de dépense comme la cantine ou l'internat, etc.
- Certaines formations ont donné des informations incomplètes, ce qui pouvait parfois s'avérer gênant. Notamment, quelques formations privées indiquent un tarif de base qui doit être complété en fonction des revenus des parents, et renvoient le candidat vers leur site internet pour des informations complémentaires : or, il s'avère que cette « part variable » dont parlent les formations en question, s'avère parfois aussi importante que la part fixe qui figure sur la plateforme Parcoursup.

Pour toutes les raisons susmentionnées, de nombreuses approximations ont été réalisées. Ainsi, les analyses qui pourraient en être tirées ne permettent de donner qu'une idée très générale du niveau réel des montants des frais de scolarité des formations présentes sur Parcoursup.

#### **Éléments associés à l'annexe 6 (« Annexe n° 6 : la valorisation des données issues d'APB 2017 à Parcoursup 2019 pour répondre aux enjeux de l'enquête »)**

Les variables susceptibles d'être utilisées par les CEV pour trier les candidats que la Cour a considéré sont détaillées dans le tableau présenté sur l'archive ouverte. Dans l'optique de l'inférence, les données sont, si nécessaire, transformées en valeurs numériques. Les valeurs manquantes sont imputées par la méthode de la médiane.

***Tableau n° 6 : Ensemble des critères calculés par la Cour pour rendre compte des pré-classements réalisés par les CEV ("algorithmes locaux") ainsi que des classements opérés par la plate-forme nationale ("algorithmes nationaux").***

<b><i>Catégorie de critère</i></b>	<b><i>Critère</i></b>	<b><i>Valeurs</i></b>	<b><i>Caractère facilement automatisable</i></b>	<b><i>Nom de la variable</i></b>	<b><i>Périmètre</i></b>	<b><i>Table(s) employée(s) et observations</i></b>
<b><i>Résultats et dossier scolaire</i></b>	Moyenne simple au bac (sans pondération des matières)	Entre 0 et 20	Oui	RES_AUTO_moyenne_simple_bac	APB + PS	i_can_epr_bac
	Moyenne générale au bac (pondération des matières variable suivant les séries)	Entre 0 et 20	Oui	RES_AUTO_moyenne_generale_bac "	APB + PS	i_can_epr_bac

<b>Catégorie de critère</b>	<b>Critère</b>	<b>Valeurs</b>	<b>Caractère facilement automatisable</b>	<b>Nom de la variable</b>	<b>Périmètre</b>	<b>Table(s) employée(s) et observations</b>
	Moyenne aux deux épreuves de français anticipé du bac	Entre 0 et 20	Oui	RES_AUTO_moyenne_francais_anticipe_bac	APB + PS	i_can_epr_bac
	Moyenne simple au contrôle continu de première et Terminale (CC)	Entre 0 et 20	Oui	RES_AUTO_moyenne_simple_CC	APB + PS	i_bul_sco
	Variabilité des notes au CC	Réel positif	Oui	RES_AUTO_variability_notes_CC	APB + PS	i_bul_sco. L'écart type est employé comme mesure de variabilité.
	Moyenne des notes d'EPS au CC	Entre 0 et 20	Oui	RES_AUTO_moyenne_EPS_CC	APB + PS	i_bul_sco
	Moyenne des notes de LV1 au CC	Entre 0 et 20	Oui	RES_AUTO_moyenne_LV1_CC	APB + PS	i_bul_sco
	Moyenne des notes d'histoire-géographie au CC	Entre 0 et 20	Oui	RES_AUTO_moyenne_histoire_geographie_CC	APB + PS	i_bul_sco
	Moyenne des notes de LV2 au CC	Entre 0 et 20	Oui	RES_AUTO_moyenne_LV2_CC	APB + PS	i_bul_sco
	Moyenne des notes de mathématiques au CC	Entre 0 et 20	Oui	RES_AUTO_moyenne_mathematiques_CC	APB + PS	i_bul_sco
	Moyenne des notes de philosophie au CC	Entre 0 et 20	Oui	RES_AUTO_moyenne_philosophie_CC	APB + PS	i_bul_sco
	Moyenne des notes de français au CC	Entre 0 et 20	Oui	RES_AUTO_moyenne_francais_CC	APB + PS	i_bul_sco
	Moyenne des notes de physique-chimie au CC	Entre 0 et 20	Oui	RES_AUTO_moyenne_physique_chimie_CC	APB + PS	i_bul_sco
	Moyenne des notes de SVT au CC	Entre 0 et 20	Oui	RES_AUTO_moyenne_SVT_CC	APB + PS	i_bul_sco

<b>Catégorie de critère</b>	<b>Critère</b>	<b>Valeurs</b>	<b>Caractère facilement automatisable</b>	<b>Nom de la variable</b>	<b>Périmètre</b>	<b>Table(s) employée(s) et observations</b>
	Moyenne des notes en sciences au CC	Entre 0 et 20	Oui	RES_AUTO_moyenne_sciences_CC	APB + PS	i_bul_sco
	Moyenne des notes en SES au CC	Entre 0 et 20	Oui	RES_AUTO_moyenne_SES_CC	APB + PS	i_bul_sco
	Moyenne des notes de mathématiques de spécialité au CC	Entre 0 et 20	Oui	RES_AUTO_moyenne_math_spe_CC	APB + PS	i_bul_sco
	Moyenne des notes d'économie-droit au CC	Entre 0 et 20	Oui	RES_AUTO_moyenne_economie_droit_CC	APB + PS	i_bul_sco
	Moyenne des notes en enseignement technologique au CC	Entre 0 et 20	Oui	RES_AUTO_moyenne_enseignement_techno_pro_CC	APB + PS	i_bul_sco
	Moyenne des notes en prévention santé environnement au CC	Entre 0 et 20	Oui	RES_AUTO_moyenne_prevention_sante_environnement_CC	APB + PS	i_bul_sco
	Effectif moyen du candidat dans sa classe de Terminale (composante de la fiche « Avenir »)	Entier positif	Oui	RES_AUTO_effectif_moyen_classe	APB + PS	i_ael_prf <sup>1</sup>
	Position relative moyenne du candidat dans sa classe de Terminale (composante de la fiche « Avenir »)	Entre 0 (dernier de classe) et 1 (1er de classe)	Oui	RES_AUTO_position_relative_moyenne_eleve	APB + PS	i_ael_prf Calcul : $[1 - (\text{rang} / \text{effectif})]$ .
	Moyenne des quatre appréciations normées du conseil de classe et de l'équipe pédagogique, coordonnée par le professeur principal,	Entre 0 (appréciation la plus faible) et 20 (appréciation la plus élevée)	Oui (variables catégorielles normées et ordonnées)	RES_AUTO_Avenir_PP_moyenne	PS	i_ael_pp <sup>3</sup> Calcul : chaque appréciation contient quatre modalités

<sup>1</sup> Le terme « ael » indiqué dans la dénomination de la table de Parcoursup signifie « appréciations en ligne ». La fiche « Avenir » est générée automatiquement sur la base des éléments renseignés par les professeurs (« prf ») et les professeurs principaux (« pp ») de l'élève.

<sup>3</sup> Cf. note *supra* relative aux « ael ».



Catégorie de critère	Critère	Valeurs	Caractère facilement automatisable	Nom de la variable	Périmètre	Table(s) employée(s) et observations
	dans le cadre de la fiche « Avenir » <sup>2</sup> .					ordonnées. Elles peuvent donc être converties sur une échelle [0, 20] par le biais d'une transformation affine.
Motivation	Taux de fautes d'orthographe dans le projet de formation motivé (PFM)	Entre 0 (0 % de fautes) et 1 (100 % de fautes)	Non (pas sans capacités d'analyses textuelles avancées)	PFM_MAN_orthographe	APB + PS	Cette variable dépend d'un couple (candidat, formation), et pas uniquement du candidat
	Richesse du vocabulaire employé dans le PFM	Entre 0 (PFM pauvre) et 1 (PFM riche)	Non (pas sans capacités d'analyses textuelles avancées)	PFM_MAN_richesse	APB + PS	Formule : logarithme du nombre de mots unique rapporté au logarithme du nombre total de mots. Cette variable dépend d'un couple (candidat, formation), et pas uniquement du candidat
	Adéquation de premier niveau entre le PFM et la formation visée par le candidat (voir détails dans l'encadré <i>infra</i> )	0 (aucune adéquation) ou 1 (adéquation totale)	Non (pas sans capacités d'analyses textuelles avancées et un cluster de calcul)	PFM_MAN_adequation_premier_niveau	APB + PS	Formule : variable binaire qui prend la valeur 1 lorsque la classe thématique de la formation est la même que celle prédite pour le projet de formation motivé. Cette variable dépend d'un couple (candidat, formation), et pas uniquement du candidat.

<sup>2</sup> Les quatre variables de la fiche Avenir sont : « i\_pp\_met\_tra » (méthodes de travail), « i\_pp\_aut » (autonomie), « i\_pp\_eei » (engagement, esprit d'initiative au lycée et en extrascolaire) et « i\_pp\_cai » (capacité à s'investir dans le travail). Chacune de ces variables comporte quatre niveaux ordonnés : « très satisfaisante », « satisfaisante », « assez satisfaisante » et peu « démontrée ». En 2019, la variable « i\_pp\_eei », qui porte toujours le même nom, a été modifiée pour refléter l'engagement citoyen en deux niveaux : « oui » ou « non ». Les corrélations de rang entre ces variables étant très élevées (entre + 0.56 et + 0.87 en 2018), prendre la moyenne permet de regrouper l'information contenue dans ces variables sans la dégrader.

<b>Catégorie de critère</b>	<b>Critère</b>	<b>Valeurs</b>	<b>Caractère facilement automatisable</b>	<b>Nom de la variable</b>	<b>Périmètre</b>	<b>Table(s) employée(s) et observations</b>
						Calcul pour l'année 2018 seulement (capacité de calcul limitée).
<i>Le profil et la filière du candidat</i>	Lycéen scolarisé dans le secondaire	0 (non) ou 1 (oui)	Oui	PROFIL_AUT_O_neo_entrant_sup	APB + PS	g_can et g_pro
	Candidat – étudiant dans le supérieur en réorientation	0 (non) ou 1 (oui)	Oui	PROFIL_AUT_O_reo_sup	APB + PS	g_can et g_pro
	Candidat en reprise d'études	0 (non) ou 1 (oui)	Oui	PROFIL_AUT_O_reprise_etudes	APB + PS	g_can et g_pro
	Nombre de redoublements de la première et/ou de la terminale pour les candidats scolarisés en Terminale en année d'affectation courante	0 (pas de redoublement) , 1 (un redoublement) , 2 (deux redoublements)	Non (pas sur la base des données accessibles aux CEV)	PROFIL_MAN_nb_redoublements	APB + PS	g_can, i_can_sco et i_niv_etd
	Candidat en année préparatoire au supérieur en année d'affectation courante	0 (non) ou 1 (oui)	Oui	PROFIL_AUT_O_annee_preparatoire_au_sup	APB + PS	g_can, i_can_sco et i_niv_etd
	Voie générale, série S	0 (non) ou 1 (oui)	Oui	PROFIL_AUT_O_bac_S	APB + PS	g_can, i_cla, i_typ_cla
	Voie générale, série ES	0 (non) ou 1 (oui)	Oui	PROFIL_AUT_O_bac_ES	APB + PS	g_can, i_cla, i_typ_cla
	Voie générale, série L	0 (non) ou 1 (oui)	Oui	PROFIL_AUT_O_bac_L	APB + PS	g_can, i_cla, i_typ_cla
	Voie technologique	0 (non) ou 1 (oui)	Oui	PROFIL_AUT_O_bac_techno	APB + PS	g_can, i_cla, i_typ_cla
	Voie professionnelle	0 (non) ou 1 (oui)	Oui	PROFIL_AUT_O_bac_pro	APB + PS	g_can, i_cla, i_typ_cla
<i>Les critères sociaux</i>	Candidat boursier du secondaire	0 (non) ou 1 (oui)	Oui	SOC_AUTO_boursier_secondaire	APB + PS	g_can
	Candidat boursier du supérieur	0 (non) ou 1 (oui)	Oui	SOC_AUTO_boursier_supérieur	APB + PS	g_can
	Catégorie socio-professionnelle (CSP) du candidat d'après les CSP de ses	0 (les deux RL appartiennent à des CSP inférieures) à 1	Non (pas sur la base des données)	SOC_MAN_csp_insee_val	APB + PS	g_can, i_csp.  Potentiel indicateur de discrimination

<b>Catégorie de critère</b>	<b>Critère</b>	<b>Valeurs</b>	<b>Caractère facilement automatisable</b>	<b>Nom de la variable</b>	<b>Périmètre</b>	<b>Table(s) employée(s) et observations</b>
	représentants légaux (RL), transformée sur une échelle 0 à 1 <sup>4</sup> . La CSP s'entend ici au sens des PCS de l'INSEE <sup>5</sup> .	(les deux RL appartiennent à des CSP supérieures)	accessibles aux CEV)			positive de candidats favorisés.
	Variable indirecte <sup>6</sup> (« proxy data ») reflétant le quotient familial du foyer fiscal du candidat	Réel positif	Non (pas sur la base des données accessibles aux CEV)	SOC_MAN_proxy_quotient_familial	APB + PS	g_can.  Potentiel indicateur de discrimination positive de candidats favorisés.
	Echelon prédit dans le cadre de l'avis conditionnel d'attribution des bourses (« ACB ») du supérieur proposées par le CNOUS. Le calcul est inspiré de la fonction du SCN intitulée « calc_g_cn_flg_simu_brs.sql » <sup>7</sup> .	Entier prenant les valeurs suivantes : -1 (pas de bourse), puis de 0 (échelon 0), à 7 (échelon 7)	Non (pas sur la base des données accessibles aux CEV)	SOC_MAN_acb_ech	APB + PS	g_can, g_acb.
	Le candidat a-t-il déclaré être en situation de	0 (non) ou 1 (oui)	Oui	SOC_AUTO_handicap	APB+PS	g_can, i_ins

<sup>4</sup> La CSP de chaque RL (43 niveaux non-ordonnés) est transformée en variable numérique avec trois valeurs possibles : 0 (CSP inférieures), 0,5 (CSP intermédiaires) et 1 (CSP supérieures). Il est supposé ici que la CSP du candidat correspond à la moyenne des CSP des deux RL.

<sup>5</sup> Cf. la nomenclature des professions et catégories socioprofessionnelles (PCS) de l'INSEE (<https://www.insee.fr/fr/metadonnees/pcs2003/categorieSocioprofessionnelleAgregree/8?champRecherche=true>).

<sup>6</sup> Dans la mesure où la structure du foyer (représentants légaux mariés, pacés, en concubinage, séparés par exemple) n'est pas déclarée par les candidats sur APB et Parcoursup, cette variable ne reflète qu'incomplètement le quotient familial (<https://www.economie.gouv.fr/particuliers/quotient-familial>). Toutefois ces deux variables sont très probablement fortement corrélées, on parle donc de donnée indirecte (« proxy data » en anglais) dans le vocabulaire consacré.

<sup>7</sup> Après examen de la fonction du SCN, la Cour remarque que les principales hypothèses de calcul au titre de l'année 2019 ne reflètent pas avec une parfaite exactitudes les critères du CROUS (voir : <https://www.aide-sociale.fr/simulation-montant-bourse-crous/>) : d'une part ces bourses du supérieur s'échelonent en effet sur 8 niveaux et non pas sur 9 comme cela est indiqué dans la table g\_acb et calculé dans la fonction « calc\_g\_cn\_flg\_simu\_brs.sql » ; d'autre part le calcul des points à charge ne prend pas en compte le critère d'éloignement du domicile.

<b>Catégorie de critère</b>	<b>Critère</b>	<b>Valeurs</b>	<b>Caractère facilement automatisable</b>	<b>Nom de la variable</b>	<b>Périmètre</b>	<b>Table(s) employée(s) et observations</b>
	handicap pour au moins l'une de ses candidatures ?					
<i>Les critères géographiques</i>	Le candidat est-il considéré dans le secteur de la formation à laquelle il candidate ?	0 (non) ou 1 (oui)	Oui	GEO_AUTO_intra_sec	PS	g_can et i_ins.  Cette variable dépend d'un couple (candidat, formation), et pas uniquement du candidat.
	L'académie du bac du candidat correspond-elle à l'académie de la formation à laquelle il candidate ?	0 (non) ou 1 (oui)	Oui	GEO_AUTO_intra_aca	APB + PS	g_can, g_aca.  Cette variable dépend d'un couple (candidat, formation), et pas uniquement du candidat.
	Le département de résidence du candidat correspond-il au département de la formation à laquelle il candidate ?	0 (non) ou 1 (oui)	Oui	GEO_AUTO_intra_dpt	APB + PS	g_can, g_aca. Cette variable dépend d'un couple (candidat, formation), et pas uniquement du candidat
	Le candidat habite-t-il dans une grande ville ?	0 (non) ou 1 (oui)	Plutôt non (pas sans opérer des tris et des regroupements sur les données)	GEO_MAN_grande_ville	APB + PS	g_can, croisée avec les données de l'INSEE relatives aux catégories d'agglomération
	Le candidat habite-t-il dans une petite à moyenne ville ?	0 (non) ou 1 (oui)	Plutôt non (pas sans opérer des tris et des regroupements sur les données)	GEO_MAN_petite_moyenne_ville	APB + PS	g_can, croisée avec les données de l'INSEE relatives aux catégories d'agglomération
	Le candidat habite-t-il en zone rurale ?	0 (non) ou 1 (oui)	Plutôt non (pas sans opérer des tris et des regroupements sur les données)	GEO_MAN_rural	APB + PS	g_can, croisée avec les données de l'INSEE relatives aux catégories d'agglomération
	Le candidat habite-t-il en Outre-Mer ?	0 (non) ou 1 (oui)	Plutôt non (pas sans opérer des tris et des	GEO_MAN_Outre_Mer	APB + PS	g_can, g_aca

Catégorie de critère	Critère	Valeurs	Caractère facilement automatisable	Nom de la variable	Périmètre	Table(s) employée(s) et observations
			regroupements sur les données)			
L'identité du candidat	Le lycée d'origine du candidat représenté par la fraction de lauréats d'un bac mention « très bien » (TB) dans ce lycée l'année précédente (effet de « réputation »)	0 (0 % de mentions TB) à 1 (100 % de mentions TB)	Plutôt non (pas sans opérer des tris et des regroupements sur les données)	ID_MAN_lycee_origine_TB	APB + PS	Pour éviter les biais, cet indice a été calculé sur la bases des données de l'année précédente : g_can, g_pro, i_bul_sco et i_can_epr_bac.
	Le lycée d'origine du candidat représenté par un indice reflétant l'écart de notation entre les résultats au bac et les résultats du contrôle continu en première et terminale des élèves dans ce lycée l'année précédente (effet de « réputation »).	-1 (sur-notation extrême) à 1 (sous-notation extrême)	Plutôt non (pas sans opérer des tris et des regroupements sur les données)	ID_MAN_lycee_origine_ecarts_bac_CC	APB + PS	Pour éviter les biais, cet indice a été calculé sur la bases des données de l'année précédente : g_can, g_pro, i_bul_sco et i_can_epr_bac.  Calcul par lycée : $[(\text{moyenne au bac} - \text{moyenne au CC}) / (\text{moyenne au bac} + \text{moyenne au CC})]$
	Sexe féminin	0 (garçon) ou 1 (fille)	Oui (variable catégorielle binaire connue des CEV)	ID_AUTO_sexe	APB + PS	g_can
	Nationalité française	0 (étranger) ou 1 (français)	Non (variable inconnue des CEV ; utilisée ici comme donnée indirecte)	ID_MAN_nationalite_FR	APB + PS	g_can.  Potentiel indicateur de discrimination négative fondée sur l'origine du candidat.
	Prénom <sup>8</sup>					

Source : Cour des comptes, à partir des données issues d'APB et Parcoursup. APB : Admission PostBac. PS : Parcoursup.

<sup>8</sup> L'analyse de critères patronymiques issus de la table g\_can n'a pas révélé de cas de discriminations fondés sur l'emploi de tels critères.