

Inégalités sociales et parcours universitaire. Étude microéconométrique par la régression LOGIT

Magali Jaoul-Grammare*

« À partir du moment où il existe des strates sociales, il existe aussi des disparités devant l'enseignement qui apparaissent d'autant plus marquées qu'on s'adresse à des niveaux plus élevés du système scolaire », Raymond Boudon (1973).

Les¹ sociétés développées sont soumises à de nombreuses inégalités sociales dont l'accès aux différents stades de la scolarité est l'une des plus importantes. Depuis la fin du XX^e siècle, le projet d'une école laïque et républicaine s'étant maintes fois heurté à des obstacles majeurs, le système d'enseignement fait l'objet de nombreuses controverses, tant sur le plan de son équité que sur celui de son efficacité.

L'efficacité d'un système éducatif se mesure à sa capacité de permettre à tous de réussir afin de s'insérer dans la société. Pour Sen (1993), l'éducation est un bien premier dont l'acquisition est nécessaire à l'intégration de chacun dans la société.

L'équité et l'efficacité sont souvent présentées comme les priorités de base des systèmes éducatifs de bon nombre de pays occidentaux, les gouvernements faisant référence à des principes de justice lors des réformes des modes d'organisation du système éducatif. Pour Meuret (1998), quatre de ces principes interviennent dans le domaine de l'éducation :

- *efficacité* : une plus grande attention est portée aux plus aptes dans l'optique de maximiser les performances individuelles et globales ;
- *égalité* : il s'agit de procurer à chacun un service égal² ;
- *proportionnalité* : on fait prévaloir la proportionnalité des ressources par rapport aux besoins ;
- *équité* : accorder une attention plus particulière aux individus les moins favorisés³.

Ainsi, sera considéré comme plus équitable un système éducatif qui fera moins dépendre les résultats des individus de leur origine sociale.

Face à cela, le système français doit assurer deux missions peu conciliables : d'une part, dans un but d'homogénéisation sociale, assurer la réussite de chacun ; d'autre part, permettre et légitimer une différenciation qui amènera les jeunes à des situations sociales et professionnelles inégales.

En ce sens, le système éducatif français porte la marque de son histoire : une volonté perpétuelle de promouvoir l'école Régalienne qui offre au Royaume puis à l'État les cadres dont ils ont besoin.

Toutefois, si la III^e République crée l'école du peuple, elle ne crée pas l'école unique et si les Trente Glorieuses amènent une massification de l'enseignement (Lévy-Garboua 1979), la démocratisation attendue se heurte rapidement au poids des facteurs socioéconomiques tant et si bien que le système éducatif n'est rien d'autre qu'un lieu de reproduction des différences sociales, « *les titres donnant droit à profession* ». (Bourdieu 1988, p. 26). En effet, la réussite scolaire est fortement corrélée avec l'origine sociale – notion de capital culturel. Ce dernier se définit comme l'ensemble des instruments de connaissance, d'expression et de savoir-faire transmis par la famille et qui contribuent fortement à la réussite scolaire.

Ainsi, le poids relatif de ce capital social (Coleman 1964) conduit les individus à faire des choix d'orientations scolaires, au détriment parfois de leur réussite. On assiste donc à une sorte de reproduction et de légitimation du capital culturel qui conditionne les choix stratégiques des familles.

* Magali Jaoul-Grammare, CNRS BETA, Université Louis Pasteur, 61 avenue de la Forêt noire, 67 085 Strasbourg, jaoulgrammare@cournot.u-strasbg.fr.

¹ L'auteur remercie Thierry Blayac et Claude Diebolt pour leurs remarques et commentaires qui ont permis l'élaboration de cet article. Je demeure néanmoins seule responsable de l'approche développée ici.

² Il s'agit ici du principe plus connu sous le terme « traitement égal des égaux ».

³ Le principe évoqué ici est celui du « *maximin* » de Rawls.

Force est de constater toutefois, que malgré une ouverture de l'accès à l'éducation, la réussite des enfants est toujours aussi dépendante du niveau d'éducation des parents (Hanushek & Luque 2003). Si la poursuite d'études dépend de multiples facteurs, le rôle de l'origine sociale n'est pas négligeable. En ce sens, notre recherche se situe en extension des travaux précurseurs de Freeman (1971), d'une publication de Ehrenberg & Smith (1994) et de précédentes études cliométriques en termes d'engorgement (Canals & Jaoul 2005 ; Diebolt 1994, 1997, 2001 ; Diebolt & El Murr 2003, 2004 ; Jaoul 2004ab). En nous inspirant plus particulièrement du modèle de Boudon, nous cherchons à déterminer par une modélisation logistiqu, les facteurs déterminant le parcours universitaire. Notre article est scindé en quatre parties : après un rappel des fondements théoriques, nous présentons la base de données utilisée et la méthodologie appliquée. Enfin, nous discutons les résultats issus de notre analyse économétrique.

1. Fondements théoriques

En 1973, R. Boudon développe un modèle dont la problématique de base est simple : les inégalités reflètent-elles des dispositions différentes vis-à-vis de l'école, ou bien sont-elles le résultat de décisions faites par les familles d'origines sociales différentes, à chaque étape du cursus scolaire ?

Selon Bernstein (1975), c'est la première proposition qui prévaut : l'acquisition des connaissances scolaires dépend du milieu social de l'enfant. Il rejoint ici la vision de Bourdieu pour qui les inégalités scolaires sont liées au fonctionnement de l'institution scolaire.

Il n'explique pas la réussite scolaire des enfants par leur talent mais par leur héritage culturel ; il considère que chaque individu possède des ressources essentiellement léguées par ses parents : le capital culturel (ce sont les connaissances en général, utilisables principalement à l'école), le capital économique (ensemble des ressources matérielles) et le capital social (étendue des relations sociales). Selon Bourdieu, les individus cherchent à maintenir ou à améliorer leur position sociale ; l'école est alors un instrument de reproduction sociale au service des classes dominantes et tant qu'elle ne traitera pas identiquement des enfants issus de milieux différents, elle n'aura comme seul résultat que de reproduire les inégalités sociales. Ainsi, l'école ne favorise pas l'égalité des chances mais renforce les inégalités.

Pour Boudon, la seconde thèse est la plus probable : les dispositions sociales ont un rôle mineur en comparaison à la succession d'arbitrages réalisés. Son analyse diffère de celle de Bourdieu dans le sens où il considère l'école comme neutre. Pour lui, le système scolaire est caractérisé par un ensemble de points de bifurcation tel qu'à chacun d'entre eux, il existe des stratégies individuelles, fonction de l'origine sociale.

Il considère donc que la probabilité d'entrer dans l'enseignement supérieur résulte d'une succession de probabilités de « survie » dans le milieu scolaire, à différents points de bifurcation qu'il recense au nombre de huit :

- terminer ou pas le cycle élémentaire ;
- entrer ou non dans le secondaire ;
- terminer ou pas le premier cycle du secondaire ;
- entrer ou non dans le deuxième cycle du secondaire ;
- terminer le deuxième cycle du secondaire ou abandonner en cours de cycle ;
- réussir ou non à l'examen final du secondaire ;
- entrer ou pas dans l'enseignement supérieur ;
- survivre dans l'enseignement supérieur ou abandonner en cours d'études.

À chaque étape, la probabilité de chaque éventualité dépend de l'origine sociale de l'individu ; plus un individu est issu d'une classe sociale élevée, plus la part d'individus choisissant la voie menant à des études supérieures complètes est importante. En effet, les familles comparent les coûts et les avantages à chaque point de bifurcation et l'enfant continue ses études tant que les avantages sont supérieurs au coût.

Le problème réside dans le fait que d'une part, les familles issues de milieux modestes surestiment le coût des études et sous-estiment les avantages liés à l'obtention du diplôme, et d'autre part que les familles aisées raisonnent de façon contraire.

Ceci s'explique par deux phénomènes :

- dans les classes aisées, il existe un meilleur héritage culturel et par là même, la réussite dans la filière dite « noble », y est plus probable ;

- dans ces mêmes classes, la motivation pour la filière noble est plus forte parce que ne pas pouvoir accéder à cette situation éducative est considéré comme un échec ; à la différence, la réussite dans une filière courte est perçue comme un succès dans les classes moins aisées.

Ceci conduit Boudon à poser les hypothèses suivantes :

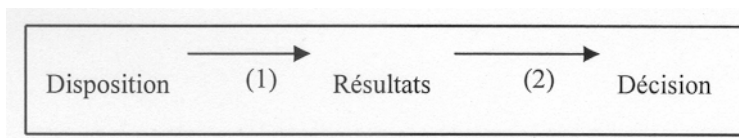
H1 : dans une génération d'élèves, certains ont plus de dispositions que d'autres pour l'école et la part d'enfants à potentiel élevé est supposée plus grande dans les classes élevées de la société ;

H2 : à origine sociale donnée, la réussite des individus ne dépend que du potentiel initial ;

H3 : à origine et potentiel initial donnés, les résultats sont les mêmes aux diverses étapes du cursus scolaire ;

H4 : à potentiel scolaire donné, la probabilité de renoncer aux études est d'autant plus faible que l'on est issu d'un milieu aisé.

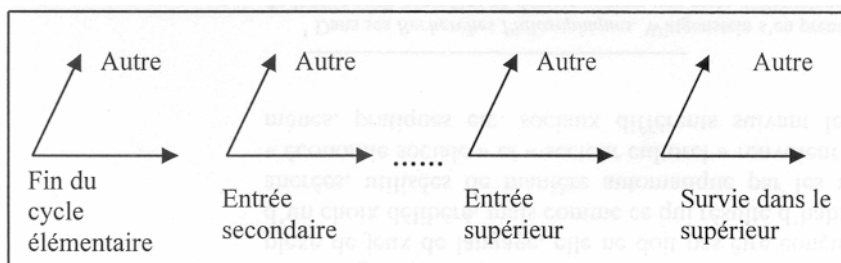
Chaque étape du cursus peut alors se schématiser comme suit (Goux et Maurin 1995, p. 96) :



Les relations (1) et (2) sont indépendantes de l'étape considérée dans le cursus mais (2) dépend de la position sociale de l'individu. Le cursus scolaire prend alors la forme d'un arbre (Figure 1) ayant huit points de bifurcation ; à chaque bifurcation, deux éventualités se présentent : survivre ou non dans le cursus scolaire.

Selon Boudon, si à la date t , un individu a la probabilité p de survivre, en généralisant, la probabilité de s'arrêter à la n ème bifurcation est égale à : $p^{n-1}(1 - p)$.

Figure 1
ARBRE REPRÉSENTANT LE CURSUS SCOLAIRE



Même en supposant que le taux de réussite ne varie pas entre les classes, Boudon parvient à des résultats où les disparités demeurent. Ce modèle comporte toutefois une critique essentielle à propos de la non-prise en compte au moment des prises de décision à chaque bifurcation, de tous les éléments pouvant jouer une influence. Le fait de supposer une indépendance entre la décision prise et ce qui se passe en aval (les problèmes de débouchés notamment) est fortement critiquable. Les conséquences du modèle rendent compte de manière relativement fidèle de l'évolution du système éducatif.

À partir de ce cadre général, nous cherchons à déterminer quelles sont les raisons de succès à un moment précis du parcours universitaire, dispositions personnelles des individus ou milieu social ?

2. Données et méthodologie

2.1. Données

Le service de pédagogie universitaire (2001) recense cinq types de critères de choix d'études supérieures :

- l'intérêt intellectuel pour un domaine particulier ;
- des atouts/faiblesses personnelles en termes de compétences ;
- les caractéristiques de la filière d'études ;
- les avantages socioéconomiques escomptés de la poursuite d'études ;
- le champ professionnel auquel donne accès le diplôme : statut social, valeurs sociales, actes professionnels.

Paradoxalement, de nombreux étudiants choisissent une filière de l'enseignement supérieur plus en fonction de ce qu'il est leur est *possible* d'y réaliser, qu'en fonction de ce qu'il leur semble *souhaitable* d'y réaliser.

Partant de là, déterminer les facteurs influençant la réussite universitaire nécessitait de disposer d'une base de données relativement bien informée quant aux variables individuelles : origine sociale, genre, scolarité... Nous utilisons pour cela la base de données Génération 98 constituée par le Céreq, suite à une enquête auprès des jeunes sortis de l'enseignement en 1998. Nous nous intéressons plus particulièrement au cursus universitaire et considérons pour cela trois seuils dans la réussite universitaire :

- poursuivre après la première année d'études ;
- poursuivre après un premier cycle ;
- poursuivre après un second cycle.

Pour chaque étape du parcours universitaire, nous cherchons ainsi à expliquer la variable y_i , telle que, $y_i = 1$ si l'étudiant franchit le seuil déterminé (1^{ère} année, DEUG, etc.)
 $y_i = 0$ s'il arrête

en fonction des divers facteurs individuels suivants :

- genre (fille, garçon) ;
- série du baccalauréat (L, ES, S) ;
- état de la scolarité jusqu'au baccalauréat (avance, normal, retard) ;
- orientation après la terminale (Deug, IUT, STS, CPGE, École ingénieur, École commerce, École de formation sanitaire et sociale, métier de la santé (médecine, pharmacie) ;
- origine (mesurée par le lieu de naissance : France, Étranger : Afrique, Amérique, Asie, Europe) ;
- origine des parents (idem) ;
- profession du père (agriculteur, artisan – commerçant – chef d'entreprise, cadre – ingénieur – profession libérale – professeur, technicien – agent de maîtrise – VRP) ;
- profession intermédiaire, employé, ouvrier ;
- profession de la mère (idem).

Notre échantillon se compose des individus ayant passé un baccalauréat général. Pour chacun des individus, lorsque certaines variables ne sont pas renseignées (absence de réponse), l'individu est supprimé de l'étude, si bien que l'échantillon retenu regroupe 15 538 individus dont tous les critères retenus ci-dessus sont parfaitement connus.

2.1. Méthodologie utilisée

De nombreuses situations de choix rencontrées dans la réalité sont incompatibles avec la théorie microéconomique classique du consommateur selon laquelle les biens sont parfaitement divisibles et où la demande est parfaitement modélisable (Deaton & Muellbauer 1980).

Considérons une population d'agents économiques rationnels faisant face à un certain nombre de biens dont la consommation est exclusive. Chaque bien procure un niveau d'utilité qui dépend des attributs des biens (Lancaster 1966) et le consommateur cherche traditionnellement à maximiser son utilité sous contrainte budgétaire, ce qui revient à décider quel bien consommer. Cependant, en raison d'une perception imparfaite de la qualité des divers biens et de la complexité à fournir un niveau exact d'utilité, on suppose que l'utilité est une fonction aléatoire (Luce 1959 ; Anderson, De Palma & Thisse 1988, 1989).

Soit un ensemble de $(m+1)$ biens dont les attributs sont regroupés dans un vecteur de variables observables, x_j , pour $j = 1 \dots m$. On note U_j l'utilité indirecte associée à la consommation du bien j (ici par exemple, l'éducation), telle que : $U_j = V(x_j) + \varepsilon_j$, où ε_j est une variable aléatoire représentant les erreurs de perception et/ou d'optimisation du consommateur, supposées indépendantes de x_j et où V est une fonction continue et déterministe.

Le problème de l'étudiant/consommateur se transforme alors en univers aléatoire. Il faut donc choisir le niveau d'éducation j qui procure la plus grande utilité indirecte aléatoire : $U_j(x_j; \varepsilon_j)$. On introduit alors une variable indicatrice y_j telle que :

$$\begin{cases} y_j = 1 \text{ si le niveau } j \text{ est choisi} \Leftrightarrow U_j = \max(U_1, \dots, U_j, \dots, U_m) \\ y_j = 0 \text{ sinon} \end{cases}$$

Les individus doivent en effet souvent choisir entre des événements mutuellement exclusifs, autrement dit, ils doivent opérer un choix discret où la demande, variable endogène, prend un nombre limité de valeurs. Le cas le plus simple consiste à choisir entre deux événements, par exemple dans notre cas « poursuivre » ou « ne pas poursuivre » ses études à un moment donné du parcours universitaire.

Dans ce cas pour un individu i , on a :

$$\begin{cases} y_i = 0 \text{ si arrêt} \\ y_i = 1 \text{ si poursuite} \end{cases}$$

Sachant que la modélisation de la demande sur un continuum de possibilités, correspond à l'espérance de la quantité achetée, le comportement de choix discret d'un individu va alors se modéliser à l'aide de la probabilité de posséder le bien ou pas.

Le problème qui se pose alors si l'on veut expliquer cette demande, est que les techniques de régression habituelles sont inadaptées, la relation entre variable endogène et variables explicatives n'étant plus linéaire. L'analyse de la covariance s'avère alors impossible et les résultats de l'étude ne sont pas fiables car :

- les estimateurs des MCO seront sans biais et convergents mais inefficaces ;
- les résidus ne seront pas normaux ; la normalité serait asymptotiquement applicable si les variances des estimateurs n'étaient pas biaisées ;
- il existe un problème d'hétéroscédasticité.

Le type de modélisation le plus utilisé alors fait appel au modèle de régression logistique (Mc Fadden 1974, 1980, 1986).

Considérons le cas d'un choix d'étude ; la variable endogène peut prendre deux valeurs :

$$\begin{cases} y_i = 0 \text{ si arrêt des études} \\ y_i = 1 \text{ si poursuite des études} \end{cases}$$

Le modèle repose sur l'hypothèse que y_i dépend d'une variable notée y_i^* et appelée variable latente, qui est fonction des variables exogènes. Elle peut être assimilée à la propension marginale à poursuivre ses études.

Le modèle devient : $\begin{cases} y_i = 0 \text{ si } y_i^* < 0 \\ y_i = 1 \text{ si } y_i^* \geq 0 \end{cases}$ avec $y_i^* = \sum_i a_i x_i + \varepsilon_i$, où x_i sont les variables exogènes et ε est le résidu.

On voit ici l'inadaptation de la méthode traditionnelle puisque la régression a pour variable endogène y_i^* , non observable. Si l'on écrit le modèle sous forme probabiliste, en notant F la fonction de répartition de la loi des résidus, on obtient :

$$\begin{aligned} P(y_i = 1) &= P(y_i^* > 0) = P(X_i'A + \varepsilon_i > 0) = P(\varepsilon_i > -X_i'A) = 1 - P(\varepsilon_i < -X_i'A) \\ &= 1 - F(-X_i'A) \\ &= F(X_i'A) \end{aligned}$$

Finalement : $y_i = \begin{cases} 1 \text{ avec la probabilité } p = F(X_i'A) \\ 0 \text{ avec la probabilité } 1 - p = 1 - F(X_i'A) \end{cases}$

L'estimation se fait par la méthode du maximum de vraisemblance. Par analogie avec la loi de Bernoulli, la loi jointe des y_i est : $[F(X_i' A)]^{y_i} [1 - F(X_i' A)]^{1-y_i}$
 Les résidus sont supposés indépendants et identiquement distribués, donc les variables y_i et y_i^* le sont également.

La vraisemblance s'écrit alors : $L = \prod_{i=1}^N [F(X_i' A)]^{y_i} [1 - F(X_i' A)]^{1-y_i}$

On déduit : $\log L = \sum_{i=1}^N \{y_i \times \log (F(X_i' A)) + (1 - y_i) \times \log (1 - F(X_i' A))\}$ ⁴

L'estimateur \hat{A} se déduit en maximisant cette fonction : $\hat{A} = \arg \text{Max}_A (\log L)$

Le modèle que nous estimons ici, est tel que F soit une fonction logistique : $F(h) = \frac{\exp(h)}{1 + \exp(h)}$, $\forall h \in \mathfrak{R}$; $F \in [0 ; 1]$.

3. Résultats

Le modèle général estimé est de la forme :

$P(Y_i = \text{poursuite}) = F(a_0 + a_1 \cdot \text{Sexe} + a_2 \cdot \text{Série} + a_3 \cdot \text{Scolarité} + a_4 \cdot \text{Classe après } T^{\text{le}} + a_5 \cdot \text{CSP Père} + a_6 \cdot \text{CSP Mère} + a_7 \cdot \text{Origine} + a_8 \cdot \text{Origine Père} + a_9 \cdot \text{Origine Mère})$

3.1. Estimation du modèle de poursuite d'étude après une année universitaire

On estime P($Y_i = \text{poursuivre après la première année universitaire}$). L'estimation porte sur 15 538 individus.

Le modèle montre que la poursuite d'études après la première année d'enseignement supérieur est influencée négativement par le sexe de l'individu et son origine. Les autres variables (Baccalauréat, État de la scolarité, Orientation après la Terminale, Origine et CSP des parents) jouent positivement ; il semblerait donc que dès le début du parcours dans l'enseignement supérieur, les variables sociales aient un rôle non négligeable dans la poursuite des études. De plus, la poursuite d'études apparaît comme fonction positive des arbitrages passés de l'individu (baccalauréat, scolarité, orientation). Le test de Wald montre que le fait de poursuivre ses études après la première année universitaire est influencé par toutes les variables.

En ce qui concerne la qualité générale du modèle, même s'il semble correct du point de vue du test de vraisemblance (il indique que $-2LL < X^2(N-p)$, ce qui signifie qu'on accepte l'hypothèse nulle selon laquelle la vraisemblance tend vers 1), on peut noter la faiblesse du PR^2 qui n'est égal qu'à 0,075. Enfin, le test des ratios de vraisemblance montre clairement que le modèle total est meilleur que les modèles restreints.

Toutefois, la faiblesse du PR^2 laisse supposer qu'en début de parcours dans l'enseignement supérieur, les facteurs personnels n'expliquent que très peu le choix d'arrêter ou pas ses études.

3.2. Estimation du modèle de poursuite d'études après un premier cycle universitaire

On estime P($Y_i = \text{poursuivre après un premier cycle universitaire}$). Les individus concernés ne sont plus qu'au nombre de 13 446.

L'estimation du modèle sur l'ensemble des variables montre que si le pouvoir explicatif demeure faible ($PR^2 = 0,074$), le rôle des variables a changé : seule la série du baccalauréat et l'origine des parents ont une influence négative sur la poursuite des études. Toutes les autres variables ont tendance à pousser à la poursuite des études après la deuxième année universitaire. Le test des ratios de vraisemblance indiquant que le modèle général est meilleur que les modèles restreints. De manière générale, le pouvoir explicatif des variables varie peu : le PR^2 est égal à 0,07. Ainsi, les facteurs sociaux influenceraient plus le choix des

⁴ Log L est appelée la log-vraisemblance.

⁵ Nous prenons ici le PR^2 au sens de Mc Fadden.

études au niveau du second cycle qu'auparavant. Vérifions cette hypothèse par l'estimation du dernier modèle.

3.3. Estimation du modèle de poursuite d'études après un second cycle universitaire

Ce dernier échantillon de 8 962 individus permet d'estimer $P(Y_i = \text{poursuivre après un second cycle universitaire})$. L'estimation du modèle montre un pouvoir explicatif plus grand que dans les modèles précédents : le $PR^2 = 0,184$. Le test des ratios de vraisemblance laisse supposer que l'origine de l'individu pourrait être ôtée de l'analyse. Le modèle restreint estimé montre que seule l'origine des parents est un facteur d'arrêt des études.

3.4. Calcul des probabilités de réussite et retour sur les hypothèses de Boudon

Dans son modèle, Boudon émet de nombreuses hypothèses quant aux divers facteurs d'influence de poursuite d'études. Nous nous intéressons plus particulièrement aux hypothèses H2 et H4 selon lesquelles :

- H2 : à origine sociale donnée, la réussite des individus ne dépend que du potentiel initial ;
- H4 : à potentiel scolaire donné, la probabilité de renoncer aux études est d'autant plus faible que l'on est issu d'un milieu aisé.

Dans le but de les vérifier, nous calculons pour les modèles 1 et 3, les probabilités de réussite des individus, en fixant les paramètres donnés.

Pour l'hypothèse H2, nous prenons l'origine sociale de référence du modèle, c'est-à-dire un individu français dont les parents sont français et qui appartiennent à la CSP « cadre ». Nous analysons plus particulièrement le cas des étudiants ayant choisi la voie universitaire. Ainsi, selon l'hypothèse H2 de Boudon, la réussite ne va dépendre que du potentiel initial de chaque individu, que nous évaluons par le baccalauréat détenu, l'état de la scolarité et le sexe.

Pour l'hypothèse H4, nous procédons en deux temps : nous prenons tout d'abord le potentiel et la nationalité comme référence, à savoir : Garçon, baccalauréat S, Scolarité avancée, Individu français de parents français et nous optons pour l'orientation universitaire (DEUG). Nous étudions les diverses probabilités selon les CSP des parents. Dans un second temps, pour le même potentiel de référence, nous fixons de plus la CSP des parents (cadre) et nous étudions les probabilités de réussite selon la nationalité des étudiants et de leurs parents.

Le calcul des diverses probabilités est présenté dans les tableaux ci-après.

Tableau 1

CALCUL DES PROBABILITÉS DE POURSUITE APRÈS LA PREMIÈRE ANNÉE D'ÉTUDES À ORIGINE SOCIALE DONNÉE

Fille	Probabilité	Garçon	Probabilité
Bac L		Bac L	
Scolarité redoublement	0,774	Scolarité redoublement	0,701
Scolarité normale	0,825	Scolarité normale	0,764
Scolarité avance	0,867	Scolarité avance	0,816
Bac ES		Bac ES	
Scolarité redoublement	0,838	Scolarité redoublement	0,780
Scolarité normale	0,877	Scolarité normale	0,830
Scolarité avance	0,908	Scolarité avance	0,871
Bac S		Bac S	
Scolarité redoublement	0,887	Scolarité redoublement	0,843
Scolarité normale	0,915	Scolarité normale	0,881
Scolarité avance	0,937	Scolarité avance	0,911

Exemple de lecture : la probabilité de poursuivre ses études après la première année d'études supérieures sachant que l'individu est une fille, titulaire d'un bac S et qui a une scolarité normale est de 0,915. Pour un garçon, elle est de 0,881.

Tableau 2

CALCUL DES PROBABILITÉS DE POURSUITE EN TROISIÈME CYCLE À ORIGINE SOCIALE DONNÉE

Fille	Probabilité	Garçon	Probabilité
Bac L		Bac L	
Scolarité redoublement	0,051	Scolarité redoublement	0,036
Scolarité normale	0,075	Scolarité normale	0,052
Scolarité avance	0,108	Scolarité avance	0,076
Bac ES		Bac ES	
Scolarité redoublement	0,076	Scolarité redoublement	0,053
Scolarité normale	0,109	Scolarité normale	0,077
Scolarité avance	0,155	Scolarité avance	0,111
Bac S		Bac S	
Scolarité redoublement	0,110	Scolarité redoublement	0,078
Scolarité normale	0,156	Scolarité normale	0,113
Scolarité avance	0,217	Scolarité avance	0,160

Tableau 3

CALCUL DES PROBABILITÉS DE POURSUITE APRÈS LA PREMIÈRE ANNÉE D'ÉTUDES À POTENTIEL DONNÉ

Père Ouvrier	Probabilité	Père Agriculteur	Probabilité
Mère Ouvrier	0,828	Mère Ouvrier	0,849
Mère Employé	0,842	Mère Employé	0,861
Mère Prof. interm.	0,854	Mère Prof. interm.	0,873
Mère Agriculteur	0,866	Mère Agriculteur	0,883
Mère Artisan, CE	0,877	Mère Artisan, CE	0,893
Mère Cadre	0,887	Mère Cadre	0,902
Père Employé		Père Artisan, CE	
Mère Ouvrier	0,836	Mère Ouvrier	0,856
Mère Employé	0,848	Mère Employé	0,868
Mère Prof. interm.	0,861	Mère Prof. interm.	0,878
Mère Agriculteur	0,872	Mère Agriculteur	0,888
Mère Artisan, CE	0,882	Mère Artisan, CE	0,898
Mère Cadre	0,892	Mère Cadre	0,906
Père Profession interm.		Père Cadre	
Mère Ouvrier	0,843	Mère Ouvrier	0,862
Mère Employé	0,855	Mère Employé	0,873
Mère Prof. interm.	0,867	Mère Prof. interm.	0,884
Mère Agriculteur	0,878	Mère Agriculteur	0,893
Mère Artisan, CE	0,888	Mère Artisan, CE	0,902
Mère Cadre	0,897	Mère Cadre	0,911

Tableau 4

CALCUL DES PROBABILITÉS DE POURSUITE EN TROISIÈME CYCLE À POTENTIEL DONNÉ

Père Ouvrier	Probabilité	Père Agriculteur	Probabilité
Mère Ouvrier	0,457	Mère Ouvrier	0,565
Mère Employé	0,469	Mère Employé	0,576
Mère Prof. interm.	0,481	Mère Prof. interm.	0,588
Mère Agriculteur	0,493	Mère Agriculteur	0,599
Mère Artisan, CE	0,505	Mère Artisan, CE	0,611
Mère Cadre	0,517	Mère Cadre	0,622
Père Employé		Père Artisan, CE	
Mère Ouvrier	0,493	Mère Ouvrier	0,600
Mère Employé	0,505	Mère Employé	0,611
Mère Prof. interm.	0,517	Mère Prof. interm.	0,622
Mère Agriculteur	0,529	Mère Agriculteur	0,633
Mère Artisan, CE	0,541	Mère Artisan, CE	0,644
Mère Cadre	0,552	Mère Cadre	0,655
Père Profession interm.		Père Cadre	
Mère Ouvrier	0,529	Mère Ouvrier	0,634
Mère Employé	0,541	Mère Employé	0,511
Mère Prof. interm.	0,553	Mère Prof. interm.	0,523
Mère Agriculteur	0,564	Mère Agriculteur	0,535
Mère Artisan, CE	0,576	Mère Artisan, CE	0,547
Mère Cadre	0,588	Mère Cadre	0,559

À origine sociale donnée, les filles réussissent mieux que les garçons ce qui est paradoxal avec le fait que les rendements des études supérieures sont supérieurs pour les garçons (Canals et Jaoul 2004). De manière générale, les individus ayant un an d'avance réussissent mieux que les autres. Enfin, quel que soit le niveau d'études analysé, le baccalauréat détenu joue dans la réussite de l'individu, la série S étant la plus « rémunératrice ». La réussite selon la CSP des parents ne présente aucune surprise : les étudiants issus de milieux relativement aisés ont plus de chance de poursuivre dans l'enseignement supérieur que ceux issus de milieux plus modestes. De plus cette différence est d'autant plus grande que l'on s'intéresse aux niveaux d'études plus élevés : notons toutefois que le croisement le plus « favorisant » n'est pas le croisement Cadre/Cadre (0,559) mais le croisement Père artisan/Mère cadre (0,655).

Ces différences de réussite en fonction de la CSP des parents s'expliquent d'une part, par le coût des études dans la mesure où les parents ne peuvent financer les études que dans une certaine limite, imposée par leur milieu social ; elles trouvent d'autre part une explication dans le fait que les étudiants issus des milieux moins aisés ont souvent recours à un travail annexe pour financer leurs études, travail qui devient peu compatible avec l'allongement de la durée des études.

Enfin, les probabilités selon les nationalités⁶ montrent que c'est essentiellement la nationalité de l'individu qui joue, les étudiants d'origine européenne ayant la probabilité de réussite la plus faible, ceux d'origine française réussissant le mieux. En ce qui concerne la nationalité des parents, ce sont les individus dont le père est asiatique qui ont les plus faibles chances de réussite. En revanche, la mère française est l'origine la plus favorable à la poursuite d'étude.

Les hypothèses de Boudon semblent relativement bien vérifiées, et ce plus particulièrement en début de parcours dans l'enseignement supérieur.

Conclusion

De manière générale, les modèles estimés montrent que plus l'individu va loin dans son parcours universitaire, plus les variables sociales ont un pouvoir explicatif important : le PR² passe de 0,09 en début de parcours à 0,20 en troisième cycle. D'après nos estimations, il semble évident que l'influence sociale n'est pas négligeable dans la poursuite des études. Toutefois, les diverses estimations montrent que c'est essentiellement en fin de parcours universitaire que s'opère le rôle des facteurs sociaux, ceux-ci expliquant plus de 20 % la poursuite des études en troisième cycle contre 8 % en début de parcours universitaire. En début de parcours, les dispositions passées des individus (baccalauréat possédé, orientation choisie, état de la scolarité) sont les principaux facteurs favorisant la poursuite d'études. À ce stade, la « massification de l'enseignement supérieur » est très présente ; ce dernier est en effet ouvert quasiment à tous, et les facteurs sociaux n'ont qu'un très petit rôle. Par la suite, seuls les étudiants issus de milieux relativement aisés peuvent prétendre à poursuivre leurs études. Une explication à ce phénomène peut être amenée par l'existence de « petits boulots » de la part des étudiants afin de financer leurs études tout au long de leur parcours universitaire ; petits boulots qu'il est de plus en plus difficile d'assumer pour les étudiants au fur et à mesure qu'ils avancent dans leurs études. Ainsi, nos résultats seraient plus en faveur d'un modèle « à la Boudon » en début de parcours, suivi d'un modèle « à la Bourdieu », ensuite. Nous retrouvons ici les conclusions de Goux et Maurin (1995) pour lesquels il existerait un « modèle hybride » proche du formalisme de Boudon par sa conceptualisation, mais incorporant des mécanismes propres à Bourdieu, qui mêle « *certaines mécanismes proprement culturels et d'autres dérivant de la stratification sociale* » (Goux et Maurin 1995, p. 116).

⁶ Les tableaux de calcul ne sont pas présentés en raison de leur taille trop volumineuse.

Bibliographie

Anderson S.P., De Palma A. & Thisse J. F. (1988), « A representative Consumer Theory of the Logit Model », *International Economic Review*, 29, pp. 461-466.

Aube R. (1982), « L'influence des facteurs socioéconomiques sur la demande privée d'enseignement universitaire », *Canadian Journal of Economics*, 15,4.

Azema C. (2002), « Favoriser la réussite scolaire », *Notes d'IENA*, n° 115.

Bernstein B. (1975), « Class and pedagogies: visible and invisible », in *Class, Codes and Control*. London: Routledge and Kegan Paul, pp. 116-156.

Boudon R. (1973), *L'inégalité des chances*, Paris, Armand Colin.

Bourdieu P. (1988), « Les déterminants sociaux de la réussite scolaire », *Entretiens au collège de France*, Paris, CNDP.

Bourdieu P. & Passeron J.-C. (1964), *Les héritiers*, Paris, Les éditions de Minuit.

Canals V. & Diebolt C. (2001), « Pourquoi entrer à l'Université ? L'exemple d'une université de Lettres et sciences humaines », *International Review of Education*, 47-6, pp. 539-572.

Canals V. & Jaoul M. (2004), « Choix d'orientation et rentabilité de l'enseignement supérieur. Une application microéconométrique à partir du modèle de scolarité de Mincer », *Brussels Economic Review*, vol. 47, 3-4, pp. 449-482.

Coleman J. (1988), « Social Capital in the Creation of Human Capital », *American Journal of Sociology*, 94, pp. S95-S120.

Deaton A. & Muellbauer J. (1980), *Economics and Consumer Behavior*, Cambridge University Press.

De Maris A. (1992), « Logit Modeling : Practical Applications », *Sage University Paper Series on Quantitative Applications in the Social Science*, n° 86, Newbury Park, CA, 87 p.

De Palma A. & Thisse J. F. (1989), « Les modèles de choix discret », *Annales d'économie et de statistique*, avril-juin, pp. 151-190.

Diebolt C. (1994), *L'évolution de longue période du système éducatif allemand*. Vol. 1 : *Éducation, croissance et cycles longs*. Vol. 2 : *L'histoire quantitative de l'éducation*, Thèse de doctorat en sciences économiques, Université Montpellier I, 15 février, 1 042 p.

Diebolt C. (1997), « L'évolution de longue période du système éducatif allemand : 19^e et 20^e siècles », *Économies et sociétés* (Special Issue), Cahiers de l'ISMEA, Série AF, n° 23, pp. 1-370.

Diebolt C. (2001), « La théorie de l'engorgement », *Économie appliquée*, 54 (4), pp. 7-31.

Diebolt C., El Murr B. (2003), « A Model of Glutting. Human Capital and Labour Markets in the Long-Run », *Applied Economics Letters*, 10, 9, pp. 557-560.

Diebolt C., El Murr B. (2004), « Educational Development and Labour Markets. The Case of Higher Education in Germany, 1820-1941 », *Quality & Quantity*, 37.

Diebolt C. & El Murr B. (2004), « A Cobweb Model of Higher Education and Labour Market Dynamics », *Mimeo LAMETA*, 25 p.

Duru-Bellat M. & Kieffer A. (1999), « La démocratisation de l'enseignement "revisitée" : une mise en perspective historique et internationale », *Les Notes de l'IREDU*, juillet.

Engle R. & Mc Fadden D. (1994), *Handbook of Econometrics*, Amsterdam, North Holland.

- Ehrenberg R., Smith R. (1994), *Modern Labor Economics. Theory and Public Policy*, 5th Ed., New York, Harper Collins.
- Freeman R. (1971), *The Market for College-Trained Manpower. A Study in the Economics of Career Choice*, Harvard University Press, Cambridge.
- Freeman R. (1976), *The Overeducated American*, New York, Academic Press.
- Freeman R. (1986), « Demand for Education », in O. Ashenfelter, R. Layard (eds), *Handbook of Labor Economics*, vol. 1, Amsterdam, North-Holland, pp. 357-386.
- Galland O. & Rouault D. (1996), « Des études supérieures inégalement rentables selon les milieux sociaux », *Insee Premières*, 469, juillet.
- Girod R. (1984), *Les inégalités sociales*, Paris, Presses universitaires de France, coll. « Que-sais-je ? ».
- Gourieroux C. (1984), *Économétrie des variables qualitatives*, Paris, Economica.
- Goux D. & Maurin E. (1995), « Origine sociale et destinée scolaire », *Revue française de sociologie*, XXXVI, pp. 81-121.
- Hanushek E. A. & Luque J. A. (2003), « Efficiency and Equity in Schools around the World », *Economics of Education Review*, 20(5), pp. 481-502.
- Jaoul M. (2004a), « Enseignement supérieur et origine sociale en France : étude statistique des inégalités depuis 1965 », *International Review of Education*. Initialement paru en *Documents de travail du LAMETA*, n° 2002-19, Université Montpellier 1, 27 p.
- Jaoul M. (2004b), « Enseignement supérieur et marché du travail : étude empirique de la théorie de l'engorgement », *Économie et prévision*, 5, 166, pp. 39-57.
- Lancaster K. J. (1966), « A New Approach to Consumer Theory », *Journal of Political Economy*, 74, pp. 132-157.
- Lemelin C. (1997), « Éducation et justice sociale », *Policy Options*, juillet-août.
- Liao T. F. (1994), « Interpreting Probability Models : Logit, Probit and Other Generalized Linear Models », *Sage University Paper Series on Quantitative Applications in the Social Science*, n° 101, Thousand Oaks, 88 p.
- Luce R. D. (1959), *Individual Choice Behavior*, New York, J. Wiley and sons.
- Mc Fadden D. (1974), « Conditional Logit Analysis of Qualitative Choice Behavior », in P. Zarembka (ed.), *Frontiers in Econometrics*, New York, Academic Press, pp. 105-142.
- Mc Fadden D. (1980), « Econometrics Models of Probabilistic Choice Among Products », *Journal of Business*, 53, pp. 513-529.
- Mc Fadden D. (1986), « The Choice Theory to Market Research », *Marketing Science*, 5, pp. 275-297.
- Meuret D. (1998), « Que serait un système éducatif juste ? », in Rapport relatif au colloque, *Le pilotage des systèmes éducatifs. Une étude des mécanismes de pilotages des systèmes éducatifs au sein des États-membres de l'Union Européenne*, Liège, du 5-8 novembre 1997.
- Petit P. (1975), « Rendement de l'enseignement supérieur et origine sociale », *Revue économique*, 26, juillet, pp. 587-604.
- Racault S. (1998), « Les filières de l'excellence », *Études et synthèses*, n° 39, août.
- Sen A. (1993), *Éthique et économie*, Paris, PUF.

Service de pédagogie universitaire (SPU) (2001), « Le projet personnel de l'étudiant : un facteur de réussite ? », *Réseau*, 49.

Thelot C. & Vallet L. A. (2000), « La réduction des inégalités sociales devant l'école depuis le début du siècle », *Économie et statistique*, 334, avril.

Vallet L. A. (1988), « L'évolution de l'inégalité des chances devant l'enseignement. Un point de vue de la modélisation statistique », *Revue française de sociologie*, XXIX, pp. 395- 423.

Vandenberghe V. (2001), « Inégalité d'accès aux diplômes supérieurs et financement public de l'enseignement. Effet Matthieu, rendement fiscal du diplôme et migration asymétrique des qualifiés », *Communication aux Journées d'étude du réseau RAPPE*, 25 et 26 juin.

Annexes

Codage des variables

Code	Variable
V001	Sexe (Fille, garçon)
V004	Baccalauréat possédé (L, ES, S)
V005	Scolarité (Redoublement, Normal, Avance)
V006	Orientation après la Terminale (Sanitaire & Social, DEUG, BTS, IUT, Santé, Ecole de Commerce, Ecole d'ingénieur, CPGE)
V007	Lieu de naissance (Asie, Afrique, Amérique, Europe, France)
V008	Lieu de naissance du père
V009	Lieu de naissance de la mère
VO10	CSP du père (Ouvrier, employé, technicien & Prof. Intermédiaire, Agriculteur, Artisan & Chef d'entreprise, Cadre)
V011	CSP de la mère

Modèle de poursuite d'études après une première année universitaire

Informations sur l'ajustement du modèle

Modèle	-2 log vraisemblance	Khi-deux	degrés de liberté	Signif.
Ordonnée à l'origine uniquement	65034,141			
Final	53095,806	11938,335	9	,000

Pseudo R-deux

Cox et Snell	,064
Nagelkerke	,107
McFadden	,072

Tests des ratios de vraisemblance

Effet	-2 log-vraisemblance du modèle réduit	Khi-deux	degrés de liberté	Signif.
Ordonnée à l'origine	53184,654	88,848	1	,000
V001	53823,384	727,578	1	,000
V004	55592,175	2496,369	1	,000
V005	54124,350	1028,544	1	,000
V006	56193,364	3097,558	1	,000
V007	53116,688	20,882	1	,000
V008	53134,536	38,731	1	,000
V009	53136,555	40,749	1	,000
V010	53308,961	213,155	1	,000
V011	53630,695	534,889	1	,000

La statistique Khi-deux est la différence dans les -2 log-vraisemblances entre le modèle final et un modèle réduit. Le modèle réduit est formé en omettant un effet du modèle final. L'hypothèse est nulle si tous les paramètres de cet effet sont égaux à zéro.

Estimations des paramètres

Yi	B	Erreur std.	Wald	degrés de liberté	Signif.	Exp(B)	Intervalle de confiance 95% pour Exp(B)	
							Borne inférieure	Borne supérieure
Parcours poursuivi								
Ordonnée à l'origine	-,791	,081	95,592	1	,000			
V001	-,379	,014	728,364	1	,000	,684	,666	,704
V004	,415	,008	2476,899	1	,000	1,514	1,490	1,539
V005	,320	,010	1006,848	1	,000	1,377	1,350	1,405
V006	,248	,005	2480,527	1	,000	1,281	1,269	1,293
V007	-7,34E-02	,016	20,221	1	,000	,929	,900	,959
V008	6,020E-02	,010	39,487	1	,000	1,062	1,042	1,082
V009	6,765E-02	,011	41,478	1	,000	1,070	1,048	1,092
V010	5,221E-02	,004	212,445	1	,000	1,054	1,046	1,061
V011	9,714E-02	,004	519,162	1	,000	1,102	1,093	1,111

Modèle de poursuite d'études après un premier cycle universitaire

Informations sur l'ajustement du modèle

Modèle	-2 log vraisemblance	Khi-deux	degrés de liberté	Signif.
Ordonnée à l'origine uniquement	86549,010			
Final	72087,143	14461,867	9	,000

Pseudo R-deux

Cox et Snell	,092
Nagelkerke	,126
McFadden	,074

Tests des ratios de vraisemblance

Effet	-2 log-vraisemblance du modèle réduit	Khi-deux	degrés de liberté	Signif.
Ordonnée à l'origine	72418,510	331,367	1	,000
V001	72654,747	567,604	1	,000
V004	72270,988	183,845	1	,000
V005	76700,282	4613,139	1	,000
V006	74036,193	1949,050	1	,000
V007	72090,661	3,517	1	,061
V008	72148,587	61,444	1	,000
V009	72137,458	50,315	1	,000
V010	73646,845	1559,702	1	,000
V011	73075,509	988,366	1	,000

Estimations des paramètres

Yi	B	Erreur std.	Wald	degrés de liberté	Signif.	Exp(B)	Intervalle de confiance 95% pour Exp(B)	
							Borne inférieure	Borne supérieure
Parcours poursuivi Ordonnée à l'origine	-1,406	,074	358,083	1	,000			
V001	,291	,012	564,468	1	,000	1,337	1,306	1,370
V004	-,103	,008	182,892	1	,000	,902	,888	,915
V005	,599	,009	4419,899	1	,000	1,820	1,788	1,852
V006	,133	,003	1842,927	1	,000	1,142	1,135	1,149
V007	2,766E-02	,015	3,539	1	,060	1,028	,999	1,058
V008	-7,24E-02	,009	60,541	1	,000	,930	,913	,947
V009	-7,32E-02	,010	49,627	1	,000	,929	,911	,949
V010	,124	,003	1552,558	1	,000	1,132	1,125	1,139
V011	,110	,004	974,488	1	,000	1,116	1,109	1,124

Modèle de poursuite d'études après un second cycle universitaire (modèle global)

Informations sur l'ajustement du modèle

Modèle	-2 log vraisemblance	Khi-deux	degrés de liberté	Signif.
Ordonnée à l'origine uniquement	59187,350			
Final	35293,788	23893,562	9	,000

Pseudo R-deux

Cox et Snell	,224
Nagelkerke	,300
McFadden	,184

Tests des ratios de vraisemblance

Effet	-2 log-vraisemblance du modèle réduit	Khi-deux	degrés de liberté	Signif.
Ordonnée à l'origine	36848,492	1554,705	1	,000
V001	35799,967	506,179	1	,000
V004	40538,112	5244,324	1	,000
V005	36453,418	1159,630	1	,000
V006	40478,282	5184,494	1	,000
V007	35294,620	,833	1	,362
V008	35452,364	158,576	1	,000
V009	35318,836	25,048	1	,000
V010	36421,329	1127,542	1	,000
V011	35415,846	122,058	1	,000

Estimations des paramètres

Yi	B	Erreur std.	Wald	degrés de liberté	Signif.	Exp(B)	Intervalle de confiance 95% pour Exp(B)	
							Borne inférieure	Borne supérieure
Parcours poursuivi	Ordonnée à l'origine	-3,863	,094	1703,706	1	,000		
	V001	,352	,016	507,556	1	,000	1,421	1,378 1,465
	V004	,744	,011	4870,966	1	,000	2,104	2,061 2,149
	V005	,403	,012	1141,504	1	,000	1,496	1,462 1,532
	V006	,233	,003	4904,305	1	,000	1,262	1,254 1,270
	V007	-1,65E-02	,018	,832	1	,362	,984	,949 1,019
	V008	-,143	,011	158,074	1	,000	,866	,847 ,886
	V009	-6,29E-02	,013	25,029	1	,000	,939	,916 ,962
	V010	,144	,004	1117,650	1	,000	1,155	1,145 1,165
	V011	4,774E-02	,004	122,282	1	,000	1,049	1,040 1,058

Modèle de poursuite d'études après un second cycle universitaire (modèle restreint)

Informations sur l'ajustement du modèle

Modèle	-2 log vraisemblance	Khi-deux	degrés de liberté	Signif.	Pseudo R-deux	
Ordonnée à l'origine uniquement	58239,402				Cox et Snell	,224
Final	34346,672	23892,729	8	,000	Nagelkerke	,300
					McFadden	,184

Tests des ratios de vraisemblance

Effet	-2 log-vraisemblance du modèle réduit	Khi-deux	degrés de liberté	Signif.
Ordonnée à l'origine	37497,058	3150,386	1	,000
V001	34852,221	505,549	1	,000
V004	39591,151	5244,479	1	,000
V005	35505,565	1158,892	1	,000
V006	39544,152	5197,480	1	,000
V008	34517,428	170,756	1	,000
V009	34375,661	28,988	1	,000
V010	35474,533	1127,860	1	,000
V011	34468,770	122,097	1	,000

Estimations des paramètres

Yi	B	Erreur std.	Wald	degrés de liberté	Signif.	Exp(B)	Intervalle de confiance 95% pour Exp(B)	
							Borne inférieure	Borne supérieure
Parcours poursuivi	Ordonnée à l'origine	-3,921	,070	3179,771	1	,000		
	V001	,351	,016	506,935	1	,000	1,421	1,378 1,465
	V004	,744	,011	4871,095	1	,000	2,104	2,061 2,149
	V005	,403	,012	1140,945	1	,000	1,496	1,461 1,531
	V006	,233	,003	4915,940	1	,000	1,262	1,254 1,270
	V008	-,146	,011	170,223	1	,000	,864	,846 ,884
	V009	-6,56E-02	,012	28,971	1	,000	,936	,914 ,959
	V010	,144	,004	1117,949	1	,000	1,155	1,145 1,165
	V011	4,775E-02	,004	122,321	1	,000	1,049	1,040 1,058