

**LES DETERMINANTS DE REUSSITE DES PRIMO ENTRANTS A
L'UNIVERSITE, SELON UN MODELE ECONOMETRIQUE. CAS UNIVERSITE
DE JIJEL**
**THE DETERMINANTS OF SUCCESS FOR UNIVERSITY PRIMO ENTRY,
ACCORDING TO AN ECONOMETRIC MODEL. JIJEL UNIVERSITY CASE**

***Samia Yeghni**

*Faculté des sciences économiques, commerciales et sciences de gestion. Université Mohamed Seddik
Benyahia Jijel (Algérie).*

yeghni_samia@univ-jijel.dz

Mounir Louadj

*Faculté des sciences économiques, commerciales et sciences de gestion. Université Mohamed Seddik
Benyahia Jijel (Algérie).*

m.louadj@univ-jijel.dz

Reçu le : 2021/02/12 Accepté le : 2021/12/28 Publication en ligne le : 2022/05/02

ABSTRACT : The aim of the study is to identify the factors having an impact on the success of first-time university entrants. Higher education plays a key role in the acquisition of knowledge. However, the imperfection of the link between success and university suggests that other factors are involved in the success of new entrants. The exploitation of the survey by questionnaire for a sample of 200 students, carried out at the University of Jijel, for the year 2015/2016, through the econometric analysis of the binary logistic regression model made it possible to identify, certain variables, such as students' cultural practice, year of obtaining the Bac, educational and professional level of parents, etc., which have a positive effect on the probability of succeeding in the first year, on the other hand, other variables have an influence negative on success, such as university accommodation, attendance "sometimes", being a salaried student.

Keywords : new entrants; questionnaire survey; binary logistic regression; University of Jijel.

JEL Classification : C01 ; C38 ; C53 ; A22.

RESUME : Le but de l'étude est d'identifier les facteurs ayant un impact sur la réussite des primo-entrants à l'université. L'enseignement supérieur joue un rôle clé dans l'acquisition du savoir. Mais, l'imperfection du lien entre la réussite et l'université laisse à croire que d'autres facteurs interviennent dans la réussite des primo - entrants. L'exploitation de l'enquête par questionnaire pour un échantillon de 200 étudiants, réalisé à l'université de Jijel, pour l'année 2017/2018, à travers l'analyse économétrique du modèle de régression logistique binaire, a permis d'identifier, certaines variables, comme la pratique culturelle des étudiants, année d'obtention du Bac, le niveau éducationnel et professionnel des parents, ...etc, qui agissent positivement sur la probabilité de réussir la première

*Auteur Correspondant

année, par contre, d'autres variables ont une influence négative sur la réussite, comme hébergement universitaire, la présence en cours « parfois », le fait d'être étudiant salarié.

Mots clés : primo-entrants, enquête par questionnaire, régression logistique binaire, université de Jijel.

1. INTRODUCTION

La réussite à l'université est devenue une préoccupation permanente pour les établissements de l'enseignement supérieur. Il est à noter que depuis que le taux d'échec ne cesse d'augmenter avec parallèlement une croissance démographique dans les différentes années de formation, cela devrait interpeller plusieurs interrogations. Aucune recherche n'a été faite sur l'impact de la réussite ou de l'échec de l'étudiant au niveau de l'université de Jijel. Nous avons décidé de se concentrer sur les étudiants de L1 cycle où la question de la réussite est la plus cruciale. Quant à la partie empirique, un questionnaire a été soumis aux étudiants de deuxième année au moment des travaux dirigés, en vue d'obtenir des informations du parcours antérieur (secondaire et le L1) et des caractéristiques socio individuelles. Cette expérience a été menée pour la période universitaire 2017-2018, ce qui a donné naissance à une base de données de 200 individus et pour lesquels on compte près de 24 variables qui capturent différents aspects de la vie de l'étudiant.

L'intérêt de notre étude est d'apporter des éléments de compréhension de la réussite des primo-entrants, pour leur assurer un meilleur accueil. Pour réaliser cet objectif, on s'est basé d'une part sur l'utilisation d'une base de données de l'enquête, contenant des informations sur la situation individuelle, le parcours secondaire de l'étudiant, sur le niveau éducationnel et professionnel des parents, et d'autre part on s'est basé, sur l'application d'un modèle économétrique (régression logistique binaire), afin d'estimer la probabilité conditionnelle de succès des primo-entrants.

Notre problématique sera donc : quels sont les facteurs qui influencent positivement la réussite des primo entrants à l'université ?

Pour répondre à la question, nous avons proposé l'hypothèse suivante :

A priori, la réussite dans l'enseignement supérieur semble liée au parcours antérieur de l'étudiant et au niveau éducationnel de la mère.

Notre article sera structuré de la manière suivante : Après une introduction globale de l'étude, nous exposons dans la première partie une brève approche théorique, reprenant les études les plus probantes concernant les déterminants socio démographiques, individuels, de la réussite universitaire. Dans la deuxième partie, nous présentons le modèle économétrique adapté à notre enquête par questionnaire. La troisième partie, a concerné l'application de la régression logistique binaire et son interprétation.

2. PARTIE I : APPROCHE THEORIQUE DES PRINCIPAUX DETERMINANTS DES PRIMO-ENTRANTS SUR LA REUSSITE A L'UNIVERSITE.

La réussite des études **Source spécifiée non valide.** exige non seulement la capacité de répondre efficacement aux exigences universitaires, mais aussi de surmonter le

déséquilibre provoqué par le passage d'un environnement connu (le milieu familial et l'enseignement secondaire) à un environnement de vie inconnu (le campus et les études universitaires). L'échec à l'université concerne principalement la première année après le baccalauréat dans l'enseignement supérieur. Cette année, c'est la plus difficile à passer, les difficultés se cristallisent toujours principalement autour de cette année de licence. Souvent, la notion de réussite est relative, c'est-à-dire que le degré de difficulté peut être différent, un étudiant qui échoue dans une filière peut réussir dans une autre filière. En réalité, la réussite de l'étudiant dépend d'une multitude de facteurs : des caractéristiques individuelles, du parcours scolaire et du niveau éducationnel et professionnel des parents. En France, pour avoir un taux de réussite en L1 plus important, des universités françaises se sont lancées dans la création de parcours par niveaux. En se basant sur le critère « mention du Bac » certaines universités **Source spécifiée non valide**, ont instauré des parcours pour les étudiants « passables » dans le but de proposer aux étudiants les plus fragiles, des cours « renforcés ». Après cinq semaines d'intégration, les étudiants seront évalués puis affiliés à un parcours approprié : Excellence, classique, renforcé et de réorientation.

2.1. Caractéristiques individuelles de l'étudiant :

Les caractéristiques individuelles sont définies par le genre, le mode d'hébergement, la présence au cours, la pratique des activités culturelles, d'être étudiant salarié,...etc. Selon certains chercheurs **Source spécifiée non valide**. « *ont analysé les déterminants de la réussite à l'Université mettent tous en évidence l'importance de la scolarité antérieure et de l'origine sociale. Ainsi, le type de baccalauréat, l'âge auquel il a été obtenu, le genre et l'origine sociale ont une influence déterminante sur l'accès et la réussite dans l'enseignement supérieur* ». Une autre étude a montré **Source spécifiée non valide**, ce sont davantage les filles qui accèdent à l'enseignement universitaire, et elles sont majoritaire avec le temps. Les garçons sont plus nombreux à redoubler, ils optent le plus souvent pour la réorientation ou à l'abandon de leurs études, mais ces données sont le reflet du passé scolaire. Outre le genre de l'étudiant, il y a d'autres variables liées aux caractéristiques sociodémographiques et aux conditions de vie de l'étudiant. Le mode d'hébergement de l'étudiant est considéré comme un lieu de découverte, c'est un paramètre important de la vie étudiante, il permet de créer des liens entre les étudiants, d'accéder à l'université et à ses infrastructures, etc. . C'est le logement le plus recherché par les étudiants, son avantage, il présente en effet un faible loyer. En revanche, une étude a montré **Source spécifiée non valide**, certains étudiants choisissent une filière plutôt qu'une autre en fonction de la distance du domicile familial. Le logement étudiant n'est pas seulement un toit, c'est un instrument de réussite. Ce mode de vie en cité universitaire, peut favoriser l'ambition de travail en groupe pour avoir de bon résultat. C'est la raison pour laquelle, les responsables cherchent à favoriser la qualité de vie des étudiants et leur insertion sociale. C'est sur ce mode de vie universitaire qu'il est possible de jouer pour augmenter les chances de réussite. Quant à l'activité professionnelle de l'étudiant non liées aux études, cette relation entre emploi et études, a montré selon approche empirique [1] qu'il y a seulement 18 % des étudiants qui exercent une activité non intégrée, estiment ainsi qu'elle a un impact négatif sur leurs résultats d'études. Le même auteur souligne, que les étudiants en emploi, « *...ce qui peut être le plus préjudiciable à la réussite, sont les horaires de travail, lorsqu'ils ne permettent pas aux étudiants d'assister aux enseignements* ». Pour la variable « la pratique des activités culturelles, comme le sport par exemple, tout le monde sait que de faire du sport, ça nous prédispose à faire du travail intellectuel, d'ailleurs, une étude [2] a montré que la pratique du sport n'est pas seulement une question d'amusement, c'est

également un bon moyen de favoriser la réussite scolaire. En effet, l'entraînement physique produit un impact positif sur les fonctions cognitives, mais également sur le comportement de l'individu, ce qui aura pour effet de l'aider dans sa réussite. Tous les étudiants ont besoin de faire une activité à côté de leurs études (activité intellectuelle, associative, musicale ou bien physique). Le sport permet à la fois une bonne hygiène de vie en tant que prévention de maladies et à la fois d'améliorer les conditions mentales et physiques de l'étudiant. Pratiquer du sport en étant étudiant permet de mieux se connaître et permet une meilleure estime de soi. La pratique du sport entre étudiant est un très bon moyen d'intégration sociale, elle à l'étudiant de se défouler après une journée de révisions ou de cours, de se changer les idées et de se décharger de tous ses problèmes, elle permet également d'oublier le stress et l'anxiété de la journée et de mieux dormir. De cette façon, l'étudiant améliore sa capacité d'analyse et ses réflexes. Enfin, la pratique d'un sport à un effet positif sur la réussite, comme l'on démontré certaines recherches « que l'éducation physique n'exerce aucun effet négatif sur la réussite des élèves et qu'elle offre au contraire les avantages physiques, sociaux, émotionnels et cognitifs » [3]. Les auteurs d'une étude parue en 2008 concluent qu'une augmentation du temps passé à faire du sport (jusqu'à une heure de plus quotidiennement), *même s'il empiète sur le temps réservé aux « matières traditionnelles »*, n'a pas d'effets négatifs sur les résultats des élèves [4].

2.2. Caractéristiques du parcours secondaire de l'étudiant :

Le parcours secondaire de l'étudiant concerne, le nombre d'années doublé en secondaire, année d'obtention du Bac, la mention du Bac, série du Bac,...etc. Les travaux sur les inégalités sociales **Source spécifiée non valide.** résultant du statut socio-professionnel des parents, pour l'accès à l'université et la réussite universitaire, sont le cumul d'inégalités sociales réalisées à travers le système scolaire. L'entrée dans l'enseignement supérieur est dépendante de la réussite dans le secondaire. Les différences **Source spécifiée non valide.** entre les genres en matière de réussite scolaire qui apparaissent le plus tôt, sont les retards par rapport au niveau de la classe et les redoublements, et ils sont plus courants chez les garçons. Les filles obtiennent généralement de meilleures notes avec un taux de réussite plus élevé dans les diplômes de fin d'études. Une étude empirique réalisé en France auprès des étudiants présents à l'examen du premier semestre selon différentes filières. L'objectif de l'étude est d'analyser l'impact des retards au secondaire sur la réussite à l'université, les résultats ont montré : « *que les différences de réussite à l'entrée à l'université sont avant tout liées à des différences de réussite et d'orientation dans le secondaire.* » **Source spécifiée non valide.** .

L'année d'obtention du Bac peut avoir un impact négatif sur la réussite de l'étudiant en première année, l'étude **Source spécifiée non valide.** a montré que la réussite dans l'enseignement supérieur semble très liée au baccalauréat obtenu et à son niveau de réussite. Le passé et le futur académique d'un étudiant ne sont pas indifférents à l'influence de son environnement familial, à sa catégorie sociale. La volonté de porter une attention particulière au passé scolaire amène à mettre en avant l'influence de l'itinéraire personnel des étudiants. Cette dimension permet de percevoir la réussite comme résultant en partie d'une trajectoire personnelle, des variables relèveront des éléments socio-démographiques et retraceront l'itinéraire académique des étudiants (âge, sexe, milieu familial, projet initial et passé scolaire).

2.3. Caractéristiques du niveau éducationnel et professionnel des parents :

Cette dimension concerne les caractéristiques des parents comme niveau d'éducation et activité professionnelle. Une étude sur les étudiants, a montré que chaque étudiant **Source spécifiée non valide**, est marqué par ses expériences académiques antérieures et par le milieu auquel il appartient. Les catégories sociales du père, de la mère et leurs niveaux de diplôme, en tant que marqueurs du milieu socio-économique et culturel d'origine, peuvent influencer la réussite, notamment par l'apport de ressources financières, culturelles ou par la familiarité aux études supérieures. Quant à l'étude empirique **Source spécifiée non valide**, réalisé en France a montré que le niveau d'éducation de la mère joue un rôle important dans la transmission de certaines capacités qui augmentent la probabilité de réussir, alors que l'activité professionnelle est non significative pour les deux parents. Certains chercheurs **Source spécifiée non valide**, de l'Université Laval et de l'Université de Montréal se sont intéressés à l'impact de l'emploi des parents sur la réussite scolaire. Les résultats de l'étude ont permis d'observer certains faits surprenants, que la moyenne scolaire des enfants ayant une mère qui travaille est supérieure à ceux dont la mère est à la maison.

3. PARTIE II : APPROCHE THEORIQUE DU MODELE DE REGRESSION LOGISTIQUE BINAIRE.

3.1. La définition du modèle de régression logistique binaire.

La régression logistique **Source spécifiée non valide**, est une technique prédictive. Elle vise à construire un modèle permettant de prédire / expliquer les valeurs prises par une variable cible à partir d'un ensemble de variables explicatives quantitatives ou qualitatives. Très peu d'études ont été consacrées sur l'influence des facteurs socio-économiques sur les résultats académiques des étudiants. Le principe de la méthode de régression logistique **Source spécifiée non valide**, est représenté mathématiquement par une variable aléatoire

$$y = \begin{cases} 1 & \text{si l'événement se produit} \\ 0 & \text{si l'événement ne se produit} \end{cases}$$

Cette variable a une distribution de Bernoulli et on note par $p = P(Y = 1)$, la probabilité de succès et $1-p = P(Y = 0)$. L'espérance mathématique et la variance de Y sont respectivement : $E(Y) = p$ et $\sigma^2(Y) = p(1-p)$

$(X_1, X_2, X_3, \dots, X_k)$ sont des variables explicatives qualitatives ou quantitatives. Dans ce cas le modèle de régression suppose : que la distribution des facteurs quantitatifs soient normales et que la relation entre les caractéristiques de l'étudiant et la réussite en session normale (ou en session de rattrapage ou avec crédit) est linéaire.

Nous voulons expliquer la réalisation de l'événement : $y : \{0,1\}$. $P_i = \text{Prob}(y_i = 1/x_i)$, et P_i correspond à la probabilité du risque, correspond à l'échec de l'étudiant en session normale (session rattrapage et avec crédit) en fonction de l'exposition ou non de plusieurs facteurs $(x_1, x_2, \dots, x_k)$. Ces facteurs sont les caractéristiques de l'étudiant (âge, genre, mode d'hébergement boursier, etc.). Le succès du modèle **Source spécifiée non valide**. Logit est le

calcul de l'Odds Ratio, OR ou rapport de cote, sa valeur s'interprète par rapport à la modalité de référence. Si $OR \neq 1$ cela signifie qu'il y a une association entre les deux groupes. Si $OR \geq 1$, cela signifie que le numérateur est supérieur au dénominateur, c'est à dire que la fréquence de l'évènement dans le groupe k est supérieure à celle dans la catégorie de référence. Si $OR \leq 1$ cela signifie que le numérateur est inférieur au dénominateur, c'est à dire que la fréquence de l'évènement dans le groupe k est inférieure à celle dans la catégorie de référence. Ainsi, la variable est exclue du modèle de départ, si elle n'est pas associée à la variable dépendante, $OR \text{ (brut)} < 1$ c'est à dire que la fréquence de l'évènement k est inférieure à la catégorie de référence. Si $OR = 1$ cela signifie que le numérateur est égal au dénominateur, c'est à dire que la fréquence de l'évènement dans le groupe k est égale à celle dans la catégorie de référence.

3.2. Les statistiques utilisées dans le modèle de régression logistique binaire.

Il existe plusieurs outils statistiques utilisés dans le modèle de régression logistique. Les probabilités de la statistique score en entrée, les probabilités de Wald ou statistique de rapport de vraisemblance en sortie, les probabilités, les résidus, les pseudo R², la valeur de la déviance, le logit, les résidus et standardisés, la valeur du levier, l'influence de Cook...etc. Plusieurs types de tableaux sont donc établis : Tableau des résultats du modèle avec l'ensemble des variables explicatives, tableau des résultats du modèle uniquement les variables jugées significatives.

Quant à l'ajustement du modèle de régression logistique, cela revient à déterminer les coefficients de la fonction de l'échantillon, c'est à dire chercher le modèle qui donne l'estimation la plus vraisemblable. Pour cela, nous recourons à la méthode du maximum de vraisemblance, qui vise à fournir une estimation des paramètres et qui maximise la probabilité d'obtenir les valeurs réellement observées sur l'échantillon. Cette technique s'effectue par itération de manière à se rapprocher le plus de la courbe logistique en forme de S. Ce qui revient à essayer toutes les combinaisons possibles pour maximiser le Log de vraisemblance (Log L). C'est grâce à l'utilisation des logiciels comme IBM SPSS Statistic21, EWIEWS10 et Minitab18, qui permettent de faciliter toutes les itérations. L'algorithme d'estimation du modèle converge vers la solution qui a le maximum de vraisemblance et la convergence est atteinte lorsque la dernière itération n'améliore que très marginalement la vraisemblance du modèle.

3.2.1. L'interprétation des coefficients de la régression logistique binaire.

Dans le modèle de régression logistique, plusieurs coefficients de régression apparaissent. Les valeurs numériques de ces coefficients n'ont pas d'interprétation directe. Seul, le signe du coefficient de régression permet de savoir si la probabilité de réussite est une fonction croissante ou non de la variable x. Un signe positif signifie que la probabilité de réussite augmentera avec la variable correspondante, un signe négatif signifie que la probabilité de réussite diminuera avec la variable correspondante. L'une des difficultés des probabilités est leur non linéarité, elle ne varie pas identiquement selon le niveau des variables indépendantes. La significativité de ces coefficients, possèdent une distribution khi deux. Nous avons utilisé : Le test de Wald **Source spécifiée non valide.**, le test de rapport

de vraisemblance (LRT :0 Likelihood Ratio Test) et le test du multiplicateur de Lagrange (LM). Ces tests de signification possèdent une distribution khi deux. En régression logistique l'hypothèse de nullité de chaque coefficient est testée au moyen du test de WALD, qui permet de vérifier la significativité de chaque coefficient.

Le test de Hosmer & Lemeshow est considéré comme une mesure de qualité d'ajustement. Les données sont rangées par ordre croissant des probabilités, puis partagées en 10 groupes au plus. Il existe d'autres critères de mesure de l'adéquation du modèle. La mesure classique de la qualité d'un ajustement du R^2 , mais il n'est pas particulièrement significatif dans les modèles à variables dépendantes binaires. Des mesures semblables de R^2 appelées pseudo- R^2 existent en plusieurs formes, Cox et Snell, le R^2 ajusté de Nagelkerke et le R^2 de Mac Fadden. Un autre pseudo R^2 de dénombrement **Source spécifiée non valide.** s'écrit de la manière suivante : Le R^2 de dénombrement est égal au nombre de prévisions correctes sur le nombre total d'observations.

3.2.2. L'évaluation de la qualité du modèle de prédiction.

Pour savoir, si le modèle est globalement significatif, et pour comparer plusieurs modèles candidats. Le calcul de l'aire sous la courbe ROC (Receiver Operating Characteristic) indique le pouvoir explicatif de prédiction du modèle. Cette aire de la courbe est considérée comme un outil [5] d'évaluation et de comparaison des modèles. Il permet de visualiser les performances, un seul coup d'œil doit permettre de voir le(s) modèle(s) susceptible(s) de nous intéresser. L'idée de la courbe ROC est de faire varier le « seuil » de 1 à 0, quand la valeur « ROC » est égale à l'unité, cela signifie que le modèle est performant, par contre, quand la valeur est inférieure à 0.5, cela correspond à un modèle médiocre. La courbe ROC (Receiver Operating Characteristic curve) c'est aussi une courbe représentative de la sensibilité et la spécificité. L'axe des ordonnées correspond à la sensibilité du modèle et l'axe des abscisses correspond à la spécificité. Pour mesurer globalement la capacité à effectuer correctement les sujets à leur classe respective. La courbe « ROC » visualise la qualité de la règle de classification. Dans l'analyse des résidus, la méthode utilisée est le tableau de classification, à partir duquel, nous calculons le taux de bon classement, c'est une statistique courante qui permet de déterminer le pourcentage d'individus correctement classés, il permet également de calculer le taux de mauvais classement de ce modèle.

4. PARTIE III : APPROCHE EMPIRIQUE DU MOD7LE DE REGRESSION LOGISTIQUE BINAIRE.

Dans cette dernière partie, nous allons présenter, la période d'enquête, les résultats de l'estimation des paramètres, ainsi que l'interprétation des résultats obtenus à partir des différents logiciels.

La période empirique de notre étude a démarré le 15 novembre au 15 décembre 2018, la population d'étude est constituée de l'ensemble des étudiants régulièrement inscrits en première année à la Faculté sciences économiques de l'université de Jijel au titre de l'année académique 2017-2018. La variable principale à l'étude est le succès de l'étudiant en fin d'année. Quant aux outils et méthode de collecte des données, sont collectés à partir de l'enquête sur les informations relatives aux variables de l'étude.

La variable dépendante: c'est la réussite de l'étudiant (en session normale, ou en session de rattrapage ou la réussite avec crédit). La variable dépendante est une variable dichotomique binaire, les variables indépendantes, concernent les caractéristiques individuelles de l'étudiant (genre, âge, mode d'hébergement, boursier, ...etc. Ainsi que les caractéristiques du parcours secondaire et universitaire de l'étudiant, et le niveau éducationnel et professionnel des parents. Pour le traitement et analyse des données, on s'est basée sur l'utilisation le logiciel SPSS 21, EvIEWS 10, Minitab 18.

4.1. Présentation des résultats du modèle de régression logistique binaire.

Notre étude se propose de déterminer les variables ayant un effet significatif sur la réussite en fin d'année en session normale, de rattrapage ou avec crédit. Le 1er modèle regroupe l'ensemble des caractéristiques de l'étudiant (caractéristiques individuelles, le parcours secondaire et le niveau éducationnel et professionnel des parents). Le résultat de la régression logistique binaire affiche trois variables significatives, à savoir l'activité culturelle, l'année de l'obtention du bac, la série du Bac. Ces variables ont un effet sur la réussite de l'étudiant en fin d'année, selon la session normale. Le 2ème modèle regroupe l'ensemble des variables, le résultat de la régression logistique binaire affiche sept variables significatives, à savoir : l'activité professionnelle de l'étudiant en parallèle avec ses études, l'année de doublement au secondaire, la mention du Bac, la présence en cours, le niveau éducationnel du père et de la mère. Ces variables significatives ont un effet sur la réussite de l'étudiant en fin d'année, selon la session de rattrapage. Le résultat du 3ème modèle, le résultat de la régression logistique binaire indique trois variables significatives, à savoir : l'activité professionnelle de l'étudiant en parallèle avec ses études, le niveau éducationnel du père et l'activité professionnelle de la mère. Ces variables significatives ont un effet sur la réussite avec crédit.

Tableau N°1 : Identification des variables significatives de la réussite des étudiants selon la méthode de régression logistique binaire.

Variables	Modalités de références	Modalités actives
Modèle M1 : Réussite selon la session normale		
1. Activité culturelle	Ne pas pratiquer sport	Pratiquer sport
2. Année obtention du Bac	Ailleurs	2017
3. Série du Bac	Autres	Economie
Modèle M2 : Réussite selon la session de rattrapage		
1. être salarié et étudiant	Non salarié	Salarié
2. Nombre d'année doublé au secondaire	Deux années de retards	Un an de retard
3. Mention du Bac	Passable ou sans mention	Assez bien
4. Présence en cours	Jamais	souvent
5. Niveau éducationnel père	Primaire	moyen ou secondaire plus.
6. Niveau éducationnel mère	Primaire	moyen
Modèle M3 : Réussite avec crédit		
1. être salarié et étudiant	Etudiant non salarié	étudiant salarié
2. Niveau éducationnel du père	Primaire	moyen ou secondaire.
3. Niveau professionnel de la mère	Primaire	moyen ou secondaire.

Source : base de données enquête

Tableau N°2 : Statistiques descriptives des variables dépendantes

Variables binaires	Modalités	Effectifs	Pourcentage
M1 : réussite selon la session normale(RSN)			
Oui	1	90	45%
Non	0	110	55%
M2 : réussite selon la session rattrapage (RSR)			
Oui	1	73	36.5
Non	0	127	63.5
M3 : réussite avec crédit (RAC)			
Oui	1	101	50.5
Non	0	99	49.5

Source : Résultat obtenu de SPSS.21.0

Au modèle M1, c'est-à-dire la réussite en session normale, le tableau N°3 indique 45% d'étudiants en session normale réussissent et 55% ne réussissent pas. Le modèle M2, correspond à la réussite en session de rattrapage, il y a 73 étudiants qui réussissent en session de rattrapage contre 127 qui ne réussissent pas, soit respectivement 36.5% et 63.5%. Quant au dernier modèle, c'est-à-dire la réussite avec crédit, on constate d'après le résultat du tableau, 101 étudiants réussissent avec crédit contre 99 qui ne réussissent pas, presque un étudiant sur deux réussisse. Le tableau suivant, va afficher, l'estimation de trois modèles, M1, M2, M3.

**Tableau N°3 : Paramètres significatives des modèles de réussite selon :
Session normale, de rattrapage et avec crédit.**

Source : Résultat obtenu de SPSS.21.0, Minitab18

Modalités	Coefficient	E.S.	Wald	ddl	Sig	Exp(B)	Inf Exp(B)	Sup Exp(B)
M1 : réussite selon la session normale(RSN)								
A7(1) pratique activité culturelle	0.689	0.341	4.085	1	0.04	1.992	1.021	3.888
B2(1) 2015 obtention du Bac	0.809	0.334	5.860	1	0.01	2.245	1.166	4.320
B5(1) Bac Economie	-0.936	0.349	7.210	1	0.007	0.392	0.198	0.777
M2 : réussite selon la session rattrapage(RSR)								
A9(2) Présent souvent en cours	-1.140	0.556	4.212	1	0.04	0.320	0.108	0.950
B1(1) Un an de retard au secondaire.	0.846	0.463	3.348	1	0.06	2.331	0.942	5.772
B3(1) A/Bien au Bac	-1.006	0.503	4.005	1	0.04	0.366	0.136	0.98
A10(1) Etudiant salarié	0.837	0.362	5.349	1	0.02	2.309	1.136	4.692
D1(1) Niveau éducationnel père moyen	-0.778	0.459	2.873	1	0.09	0.459	0.187	1.129
D1(2) Niveau éducationnel père Secondaire et plus.	-0.797	0.437	3.333	1	0.06	0.451	0.191	1.060
D2(1) Niveau éducationnel mère, moyen	0.886	0.399	4.933	1	0.02	2.426	1.110	5.302
M3 : réussite avec crédit (RAC)								
A10(1) Etudiant salarié	-0.621	0.334	3.464	1	0.06	0.537	0.279	1.034
D1(2) Niveau éducationnel père, secondaire et plus.	0.809	0.392	4.270	1	0.03	2.246	1.043	4.839

D4(2) Activité professionnelle de la mère, fonctionnaire.	1.057	0.463	5.225	1	0.02	2.879	1.163	7.130
--	-------	-------	-------	---	------	-------	-------	-------

4.2. Interprétation statistique des paramètres estimés par la régression logistique binaire.

Selon le résultat du tableau N°3, l'équation exponentielle du 1^{er} modèle donne une chance de 63.67% de réussir en session normale, pour le 2^{me} modèle, l'équation exponentielle, indique une chance de 24% de réussir en session de rattrapage. Quant au dernier modèle, l'équation exponentielle affiche une probabilité de 55.45% de chance de réussir avec crédit. L'interprétation des résultats repose sur le signe positif ou négatif des coefficients, qui indiquent si la variable a une influence positive ou négative sur la probabilité de réussir en session normale, session de rattrapage ou réussir avec des crédits. Puis, on appréciera la significativité des coefficients à l'aide des odds ratios.

Pour le modèle M1 : la Réussite en Session Normale (RSN), la régression logistique binaire, a affiché trois variables significatives. La variable « la pratique des activités culturelles » a un lien significatif ($\text{Sig} = 0.04 < \text{Seuil critique} = 0.05$), odds ratio est égal à 1.992, le signe du coefficient est positif. On peut conclure, qu'il y a une chance de presque 1.2 fois d'augmenter la probabilité de réussite en session normale de pratiquer les activités culturelles. La variable « année (2017) d'obtention du Bac » est significatif ($\text{Sig} = 0.01 < \text{Seuil critique} = 0.05$), odds ratio est égal à 2.245, le signe du coefficient est positif, on peut conclure, que l'année d'obtention du Bac, a une chance de 2.24 fois d'augmenter la probabilité de réussite en session normale pour l'année 2015.

La variable « série Bac » est hautement significatif ($\text{Sig} = 0.007 < \text{Seuil critique} = 0.01$), l'odds ratio est égal à 0.392, le signe du coefficient est négatif. On peut dire donc, qu'il y a un risque de 0.39 fois de diminuer la probabilité de réussite en session normale pour la série du Bac en 2^e économie.

Pour le modèle M2 : la Réussite en Session de Rattrapage (RSR), la régression logistique binaire, a affiché sept variables significatives. La variable « la présence en cours » a un lien significatif ($\text{Sig} = 0.04 < \text{Seuil critique} = 0.05$), odds ratio est égal à 0.320, le signe du coefficient est négatif. On peut conclure, qu'il y a un risque de 0.32 fois de diminuer la probabilité de réussite en session de rattrapage en assistant parfois au cours.

La variable « étudiant est salarié » est significatif ($\text{Sig} = 0.02 < \text{Seuil critique} = 0.05$), odds ratio est égal à 2.309, le signe du coefficient est positif, on peut conclure, que la chance qu'un étudiant est salarié à 2.309 fois d'augmenter la probabilité de réussite en session de rattrapage. La variable « le nombre d'année doublé en secondaire » est significatif ($\text{Sig} = 0.06 < \text{Seuil critique} = 0.10$), l'odds ratio est égal à 2.331, le signe du coefficient est positif. On peut dire donc, qu'il y a une chance de 2.331 fois d'augmenter la probabilité de réussite en session de rattrapage pour la modalité d'une année de doublement en secondaire.

La variable « mention du Bac », a un lien significatif ($\text{Sig} = 0.04 < \text{Seuil critique} = 0.05$), odds ratio est égal à 0.366, le signe du coefficient est négatif. On peut conclure, qu'il y a un risque de 0.366 fois de diminuer la probabilité de réussite en session de rattrapage pour la modalité de la mention du Bac assez bien. La variable « le niveau éducationnel du père » a un lien significatif ($\text{Sig} = 0.09 < \text{Seuil critique} = 0.10$), odds ratio est égal à 0.459, le signe du coefficient est négatif. On peut conclure, qu'il y a un risque de 0.459 fois de diminuer la probabilité de réussite en session de rattrapage lorsque le père a un niveau éducationnel moyen. La variable « le niveau éducationnel du père » a un lien significatif ($\text{Sig} = 0.06 <$

Seuil critique = 0.10), odds ratio est égal à 0.451, le signe du coefficient est négatif. On peut conclure, qu'il y a un risque de 0.451 fois de diminuer la probabilité de réussite en session de rattrapage lorsque le père a un niveau éducationnel secondaire et plus. La variable « le niveau éducationnel du père » est significatif (Sig = 0.02 < Seuil critique = 0.05), odds ratio est égal à 0.412, le signe du coefficient est positif. On peut conclure, que la chance qu'une mère ait un niveau éducationnel moyen à 0.412 fois d'augmenter la probabilité de réussite en session de rattrapage.

Pour le modèle M3 : Réussite Avec Crédit (RAC), la régression logistique binaire, a affiché trois variables significatives. La variable « étudiant est salarié » est significatif (Sig = 0.06 < Seuil critique = 0.10), odds ratio est égal à 0.537, le signe du coefficient est négatif. On peut conclure, que le risque qu'un étudiant est salarié est de 0.537 fois de diminuer la probabilité de réussite avec crédit. La variable « le niveau éducationnel du père » est significatif (Sig = 0.03 < Seuil critique = 0.05), odds ratio est égal à 2.246, le signe du coefficient est positif. On peut conclure, que le niveau éducationnel du père, pour la modalité « secondaire et plus », a une chance de 2.246 fois d'augmenter la probabilité de réussite avec crédit. La variable « activité professionnelle de la mère » est significatif (Sig = 0.02 < Seuil critique = 0.05), l'odds ratio est égal à 2.879, le signe du coefficient est positif. On peut dire, qu'il y a une chance de 2.879 fois d'augmenter la probabilité de réussite avec crédit, lorsque la mère est fonctionnaire.

Après estimation et interprétation des modèles, la question qui se pose, est de savoir, si ces trois modèles sont proches de la réalité et permettent de prédire le comportement des variables dépendantes. Pour cela, le recours à certains critères : AIC, les pseudo R² celui de Cox et Snell, de Mc Fadden, ...etc. s'avère important. Ces critères sont obtenus à partir du logiciel SPSS 19.0 et le logiciel Eviews 10.0.(voir tableau N°4).

**Tableau N°4 : La qualité d'ajustement des Modèles selon certains critères.
(Modèles : RSN- RSR –RAC)**

Critères	Modèle M1 : Réussite session normale	Modèle M2 : Réussite session de rattrapage	Modèle M3 : Réussite avec crédit
Mc Fadden R2	0.0919	0.1117	0.0439
AIC	1.4596	1.3759	1.5352
Schwarz	1.8059	1.7222	1.8815
LR(statistic)	25.3215	29.3150	12.1935
Proba LR	0.1894	0.0810	0.9092
SCR	43.7191	39.0977	46.964
Hosmer Lemeshow	5.682	9.766	11.49
ddl	8	8	8
Sig	0.68	0.28	0.175
% modèle correct	68.5	74.5	66.5
-2log vraisemblance	243.053	219.619	254.61
R2 Cox et Snell	0.149	0.193	0.107
R2 Ngarkkerke	0.199	0.264	0.143
Khi deux	32.163	42.878	22.624
ddl	29	29	29
sig	0.31	0.04	0.793

Source : résultats obtenus à partir SPSS21.0 et Eviews 10.0

Les principaux critères montrent que le meilleur modèle est le modèle de la session de rattrapage. Le modèle est validé à 10% (sig LR = 0.08 < seuil critique 10%). Le critère

AIC est faible (égal à 1.375) par rapport aux deux modèles (1.45 ; 1.53). La somme des carrés des résidus est faible (SCR = 39.09) comparativement aux deux modèles (SCR1= 43.72, SCR2= 46.96). La probabilité du khi deux est égal à 0.04 <5%, contrairement aux deux modèles non significatifs (sig = 0.31 et 0.79). Les pseudo R² comme celui de Mc Fadden, qui est un critère le plus adapté à la régression logistique binaire, sa valeur varie entre 1 (le modèle est très bon) et 0 (le modèle est trivial). Dans notre cas, le modèle M1(RSN) sa valeur est égale à 0.09, pour le deuxième modèle M2(RSR) sa valeur est égal à 0.11, quant au dernier modèle M3(RAC) sa valeur est de 0.04, donc, on peut dire que ces trois sont de faible qualité. La statistique log vraisemblance du premier modèle M1 est égal à 243.053 que l'on compare au khi deux avec 29 degré de liberté vaut 32.163, elle est inférieure à 243.093, ce qui conduit au rejet de l'hypothèse nulle (tous les coefficients sont nuls), donc il existe au moins un coefficient non nul. La statistique log vraisemblance du premier modèle M2 est égal à 219.619 que l'on compare au khi deux avec 29 degré de liberté vaut 42.87, elle est inférieure à 219.619, ce qui conduit au rejet de l'hypothèse nulle, donc il existe au moins un coefficient non nul. Quant au dernier modèle M3, la statistique log vraisemblance du premier modèle M2 est égal à 254.61 que l'on compare au khi deux avec 29 degré de liberté vaut 22.62, elle est inférieure à 254.61, ce qui conduit au rejet de l'hypothèse nulle (tous les coefficients sont nuls), donc il existe au moins un coefficient non nul. Il y a également le test de Hosmer Lemeshow qui est un test d'ajustement de la régression logistique, il indique dans quelle mesure nos données correspondent à celle du modèle. Dans notre cas le sig du modèle M1 est égal à 0.68 > 0.05, ce qui conclut que le modèle 1 est bien ajusté. Pour le modèle M2, le sig a une valeur égale à 0.28 > 0.05, le M2 est bien ajusté, de même pour le M3, le sig est égal à 0.17 > 0.05, donc on peut conclure que les trois modèles sont bien ajustés. Le modèle de régression logistique binaire a fait sortir, également le tableau de classification, ce tableau permet d'appréhender la qualité prévisionnelle du modèle sur l'échantillon, en comparant la probabilité estimée pour un individu d'être y=0 au seuil arbitraire de 50% (voir tableau N°5).

Tableau N°5 : Classification des modèles

Prévision	OUI	NON	Total%	Total effectifs
Modèle M1 RSN				
OUI	54	36	60%	90
NON	27	83	75%	110
			68.5%	
Modèle M2 RSR				
OUI	38	35	52.5%	73
NON	16	111	87.5%	127
			74.5%	
Modèle M3 RAC				
OUI	68	33	67.3%	101
NON	34	65	65.7%	99
			66.5	

Source : résultats obtenus à partir SPSS21.0

Le tableau N°5 montre parmi les 90 individus pour lesquels y = 1, le modèle M1 indique 54 individus ont une probabilité estimée de réussir en session normale RSN inférieure à 50%.

Dans 60% des cas qui réussissent en session normale sont correctement prévus. On constate également parmi les 110 individus pour lesquels $y = 0$ le modèle affiche 83cas, ont une probabilité estimée de ne pas réussir en session normale inférieur à 50%. Dans 75% des cas de ne pas réussir en session normale sont correctement prévus. Le taux d'erreur de 31.5% est considéré comme faible, par contre 68.5% que le modèle M1 est correcte.

Le modèle M2, affiche 52.5% des cas qui réussissent en session de rattrapage sont correctement prévus. On constate également parmi les 127individus pour lesquels $y = 0$, le modèle affiche 111cas ont une probabilité estimée de ne pas réussir en session de rattrapage inférieur à 50%. Dans 87.5% des cas de ne pas réussir en session de rattrapage sont correctement prévus. Le taux d'erreur de 24.5% est considéré comme faible, par contre 74.5% que le modèle M2 est correcte. Le modèle M3, montre 33 individus ont une probabilité estimée de réussir avec crédit inférieure à 50%. Dans 67.3% des cas qui réussissent avec crédit sont correctement prévus. On constate également parmi les 99 individus pour lesquels $y = 0$, le modèle affiche 65cas ont une probabilité estimée de ne pas réussir avec crédit inférieur à 50%. Dans 65.7% des cas de ne pas réussir avec crédit sont correctement prévus. Le taux d'erreur de 33.5% est considéré comme faible, par contre 66.5% que le modèle M3 est correcte.

4.3. Interprétation économique des résultats du modèle économétrique

L'objectif de l'étude est de quantifier la réussite en session normale, en session de rattrapage et la réussite avec crédit. D'après les pseudo R2, le Mc Fadden a montré que les trois modèles sont de faibles qualité, par contre le test de Hosmer Lemeshow atteste que les trois modèles sont bien ajustés. Pour le modèle M1, la variable qui influence positivement c'est l'année d'obtention du bac, puis la pratique culturelle. Pour le modèle M2, les variables qui influencent positivement sur la réussite selon la session de rattrapage, est le fait d'être étudiant/salarié, le fait d'avoir une année de retard en secondaire, le niveau éducationnel de la mère est moyen. Quant au dernier modèle, la variable qui a une influence positive sur la réussite avec crédit, vient en première position, l'activité professionnelle de la mère, puis vient en deuxième position, le niveau éducationnel du père « niveau secondaire et plus. A partir de cette analyse économétrique, il est du devoir des responsables de tenir compte de ces variables significatives, pour améliorer la chance de réussite selon la session normale, de rattrapage ou selon l'admission avec crédit.

4.4. L'interprétation de la courbe ROC

La courbe ROC a pour objectif de trouver un compromis acceptable entre une forte sensibilité et une forte spécificité. La sensibilité est la capacité à diagnostiquer la réussite de l'étudiant parmi la réussite. Par contre la spécificité est la capacité à reconnaître la non réussite de l'étudiant par la non réussite, et 1-spécificité correspond au risque de diagnostiquer la réussite de l'étudiant chez les étudiant qui n'ont pas réussi. Le tableau N°6, affiche un indice de performance de 0.724 selon la probabilité de prédiction logistique, pour le 1^{er} modèle, mais, la vraie valeur se trouve dans 95% de chance entre [0.654- 0.794. La qualité du modèle est performante, la probabilité : Sig=0.00 < pr = 0.05. Donc, on peut conclure qu'il y a une forte chance de 72% de réussir en session normale selon le modèle de régression logistique binaire.

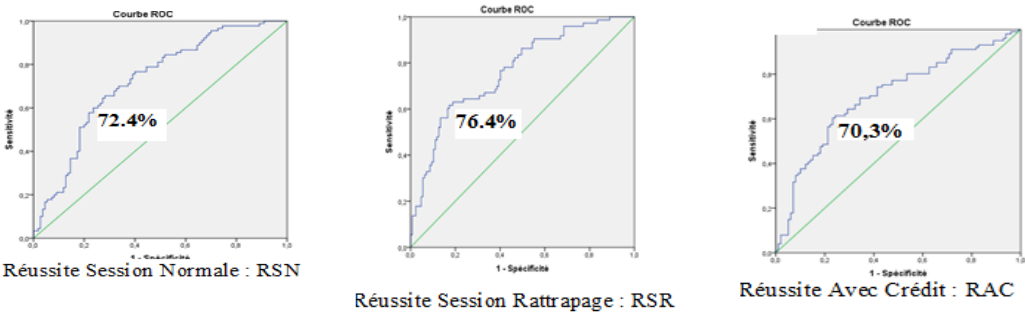
Pour le 2ème modèle, la réussite des étudiants en session de rattrapage, l'indice de performance est de 0.764, mais, la vraie valeur se trouve dans 95% de chance entre [0.696 - 0.832] , la qualité du modèle est performante, en raison de la probabilité : $\text{Sig} = 0.00 < pr = 0.05$, donc, il y a une forte chance de réussir en session de rattrapage selon le modèle logistique binaire. Quant au 3ème modèle, la probabilité de réussir avec crédit, l'échantillon affiche un indice de performance de 0.703 selon le modèle logistique, mais la vraie valeur se trouve dans 95% de chance entre [0.631-0.776], la qualité du modèle est performante, car le $\text{sig} = 0.00 < pr = 0.05$. D'après ces résultats, on peut qu'il y a une forte chance de réussir en session de rattrapage selon le modèle logistique 76.4%, puis vient en deuxième position, la chance de réussir en session normale 72.4% et finalement 70.3% est la probabilité de réussir avec crédit.

Tableau N06 : Analyse de la courbe ROC

Variables résultats tests	Zone	Erreur Standard	Sig	Intervalle de confiance 95%	
				Borne inférieure	Borne supérieure
Réussite selon la session normale RSN					
M1 : Probabilité de prédiction logistique	0.724	0.036	0.000	0.654	0.794
Réussite selon la session rattrapage RSR					
M2 : Probabilité de prédiction logistique	0.764	0.035	0.000	0.696	0.832
Réussite avec crédit RAC					
M3 : Probabilité de prédiction logistique	0.703	0.037	0.000	0.631	0.776

Source : Résultats obtenus à partir de SPSS 19.0

Figure N°1 : La représentation graphique de la courbe ROC selon le type de réussite et le pouvoir explicatif.



5. Conclusion du résultat empirique

Le résultat de l'étude empirique a fait sortir plusieurs variables ayant une influence positive sur la réussite. La variable année d'obtention du Bac est en première position du point de vue de l'importance de son influence positive sur la variable dépendante, elle augmente la chance de réussir en session normale, puis vient en 2ème position, la variable la pratique des activités culturelles. Quant au 2ème modèle, le niveau éducationnel de la mère, le nombre de retard en secondaire, et le fait d'être étudiant salarié, est en deuxième position du point de vue l'importance de son influence positive de réussir en session de rattrapage. Quant au 3ème modèle, l'étude a fait sortir deux variables agissant positivement sur la réussite avec crédit,

à savoir, l'activité professionnelle de la mère, occupe la 1^{ère} position « mère fonctionnaire » du point de vue l'importance de son influence positive sur la variable dépendante y_i , puis vient en deuxième position le niveau éducationnel du père « secondaire et plus » d'augmenter la chance de réussite avec crédit.

5. CONCLUSION

Cette étude, est loin d'identifier tous les facteurs expliquant la réussite de l'étudiant en première année de l'enseignement fondamentale, mais, elle a permis au moins de définir assez clairement certaines variables explicatives qui ont un impact positif sur la réussite (session normale, de rattrapage ou la réussite avec crédit). En effet, la modélisation économétrique de notre étude a fait apparaître certains déterminants majeurs de la réussite des primo entrants à l'université de Jijel. Les principaux résultats empiriques se sont focalisé dans : l'année d'obtention du Bac, la pratique des activités culturelles, le niveau éducationnel de la mère, le fait d'être étudiant salarié ,etc. De même, la représentation graphique de la courbe ROC a montré un fort pouvoir explicatif pour les trois modèles (supérieur à 70%). Dans, la partie théorique, nous avons constaté, que la plupart des études ont affirmé des relations positives, entre le niveau éducationnel et professionnels et le taux de réussite de l'étudiant. Notre recherche empirique, confirme bien le cadre théorique que nous avons consulté. Les résultats de l'étude ont bien confirmé notre hypothèse : la situation professionnelle de la mère influence la réussite, ce résultat a déjà été mis en évidence plusieurs fois dans la littérature. D'ailleurs, la plupart des tests économétriques que nous avons utilisés, confirment bien ces variables significatives. Nos résultats sont donc, bien conformes à ceux des études antérieures, qu'il y a une relation positive et significative entre la réussite des primo entrants et le niveau éducationnel et professionnel des parents.

Recommandations

Elargir l'échantillon d'études, afin de mieux cibler les facteurs d'impacts de réussite en 1^{ère} année ; Entreprendre des études plus large, en prenant en compte d'autres variables du vécu de l'étudiant, comme le coté social, psychologique de l'étudiant, ...etc. Ces variables peuvent peut-être améliorer les résultats académiques des primo entrants. Utiliser des modèles économétriques et faire des comparaisons avec différentes universités.

BIBLIOGRAPHIE

- [1] S. Zilloniz, «L'activité rémunérée des étudiants et ses liens avec la réussite des études. Les enseignements des enquêtes Emploi 2013-2015,» *Travail et emploi* , vol. 4, n° 1152, pp. 89-117, 2017.
- [2] «L'activité physique contribue au bon rendement scolaire,» [En ligne]. Available: <http://tictacgym.com/lactivite-physique-contribue-rendement-scolaire%E2%80%89>. [Accès le 2021 Janvier 10].

- [3] D. W. Lee Schaefer, «Quelle est l'incidence de l'éducation physique sur le bien-être des élèves et le succès scolaire?», 25 JUIN 2018. [En ligne]. Available: <https://www.edcan.ca/articles/impact-physical-education-students-well-academic-success/?lang=fr>. [Accès le 10 Janvier 2021].
- [4] T. e. Shephard, «Physical education, school physical activity, school sports and academic performance, International Journal of Behavioral Nutrition and Physical Activity 5, 1-12.,» 2008. [En ligne]. Available: <https://www.mamanpouirlavie.com/vie-scolaire/apprentissages/10943-le-lien-entre-l-exercice-physique-et-la-russite-scolaire.thtml>. [Accès le 10 Janvier 2021].
- [5] R. RAKOTOMALALA, «Receiving Operating Characteristics,» [En ligne]. Available: https://eric.univ-lyon2.fr/~ricco/cours/slides/roc_curve.pdf. [Accès le 6 Janvier 2021].
- [6] Eric. [En ligne]. Available: http://eric.univ-lyon2.fr/~ricco/cours/cours_regression_logistique.html.