



Faculté

des **sciences économiques** et de **gestion**

Université de Strasbourg

Mars 2022

Master I : Analyse et politiques économiques

Econométrie Appliquée I

Les déterminants du choix d'université des étudiants

ADUAYOM MESSAN Messan Daniel

COLIN Roudnel

DIALLO Djenabou Diari

NIANE El Makhdi Mouhamed

Sommaire

Introduction	1
Cadre Théorique	3
A. Présentation des articles de référence	3
B. Hypothèses	6
Analyse empirique.....	7
A. Analyse descriptive	7
Présentation des données	7
B. Présentation du modèle :	15
Conclusion	18

Introduction

Le choix d'une université fait partie des décisions les plus importantes qu'un étudiant est amené à prendre. Généralement, beaucoup passent le bac sans vraiment savoir vers quelle université ils voudraient se tourner pour leur formation.

Ce choix peut s'avérer d'autant plus complexe du fait de l'existence de différents types d'établissements avec des caractéristiques qui peuvent parfois être similaires, mais aussi avec des avantages en termes de perspectives d'emploi et de salaire pouvant varier d'un établissement à un autre. Le choix d'un établissement implique donc une longue réflexion car il est déterminant pour l'avenir.

Nous retrouvons cette idée à travers les propos de Diebolt (2001) « La répartition des étudiants dans les différentes universités se fait en fonction de nombreux avantages comparatifs, notamment en termes de salaires espérés et de travail disponible dans les secteurs professionnels correspondants. »

Plusieurs facteurs peuvent être pris en compte en faisant ce choix comme l'histoire familiale, les préférences (dans quelle filière l'individu veut poursuivre ses études), la qualité de la formation, le niveau d'insertion des étudiants sortant de l'établissement et leur salaire moyen mais aussi la localisation de l'établissement et le niveau de chômage dans les filières proposées par l'établissement. Notons de plus qu'avant le choix de n'importe quel établissement, une étape devra d'abord au préalable être franchie : à savoir celle du choix de la filière d'étude. Ce dernier ne peut donc se faire sur la base du hasard.

« Selon les approches en termes de capital humain, la poursuite d'études correspond à une logique d'accumulation de capital en vue d'un échange marchand sur le marché du travail » (Schultz, 1961 ; Becker, 1964 et Mincer, 1958, 1974)

De plus on peut ajouter qu'en 2001, le service de pédagogie universitaire de France a recensé cinq critères de choix d'une filière d'étude supérieurs : l'intérêt intellectuel pour un domaine, les atouts/faiblesses personnelles en termes de compétences, les caractéristiques de la filière d'étude, les avantages socio-économiques escomptés et le champ professionnel auquel donne accès le diplôme.

Comment s'effectue donc le choix d'une université ? est-ce arbitrairement ? Certainement pas. Une étape longue et difficile pour les futurs étudiants pourtant nécessaire et obligatoire. Sa difficulté réside du manque d'information dont disposent les jeunes lycéens que ça soit sur la formation en tant que telle et/ou sur leur contenu pédagogique en général. Ce qui fait que le choix de l'université ne peut guère être basé majoritairement sur le contenu. Une étude menée aux Etats Unis par Wiswall et Zafar en 2014 montre d'une part que dans le choix de filière à

l'université de New York, les étudiants n'estiment pas parfaitement les salaires associés aux filières et qu'en moyenne ils ont plutôt tendance à le sous-estimer. D'autre part, une partie des étudiants auraient fait un choix d'orientation différent et donc auraient choisi une université différente s'ils disposaient de toutes les informations sur les perspectives de carrières pour chaque filière.

Tout au long de notre étude, nous essaierons de répondre à la question à savoir : Quels sont les déterminants du choix des universités pour les étudiants ? Ainsi nous confirmeront ou infirmeront des hypothèses. Mais nous ne nous limiterons pas juste à cela, au risque de n'effleurer qu'une partie de la problématique. Nous mettrons donc en relief les facteurs qui agissent sur le choix d'une université par un étudiant quelconque et de quelle manière.

Afin de rendre notre étude cohérente, nous nous appuierons sur divers travaux et articles notamment :

- Stratégie des étudiants, choix d'orientation et problèmes d'engorgement : une analyse quasi-expérimentale de Magali Jaoul -Grammare, CNRS-BETA, Strasbourg ,2008
- Les déterminants du choix du domaine d'études universitaires : Une revue de la littérature et identification de pistes d'interventions de Claude MONMARQUETTE , CIRANO et Université de Montréal, Juin 2020

Par ailleurs, le classement Shanghai sera aussi une des ressources annexes sur lequel on s'appuiera. Ce classement répertorie les meilleures universités du monde sur la base de divers critères très pertinents tels que le nombre de prix Nobel et de médailles Fields parmi les anciens élèves et parmi les chercheurs, le nombre de chercheurs les plus cités dans leurs disciplines pendant les dix dernières années, les articles publiés dans les revues Nature et Science pendant les cinq dernières années, les articles indexés dans Science Citation Index, et Social Sciences Citation Index etc ... Un tel choix est poussé par le travail que nous devons faire.

En plus d'une étude globale sur le sujet, nous chercherons à écrire un modèle présentant le choix rationnel des universités fait par les étudiants afin de comprendre leur choix et comment différentes variables interviennent dans cet exercice ardu qu'il revient de faire à chaque nouvel étudiant pour s'assurer le meilleur choix.

Ainsi, nous présenterons dans une première partie les hypothèses sur lesquelles notre étude se base, comme dit précédemment nous disposons de nombreux travaux qui abordent notre problématique mais aussi nous présenterons les articles cités plus haut qui nous ont servis de point de départ. En seconde partie nous présenterons les différentes variables et les raisons du choix de ces variables pour notre étude. Ensuite il s'agira de réaliser tout d'abord une analyse descriptive pour comprendre la variation de nos variables, ainsi saisir leur portée et par la suite

une analyse économétrique en estimant notre modèle, l'interpréter et éventuellement le corriger. Enfin nous ferons une analyse globale de notre étude tout mentionnant les avantages, les limites mais aussi des recommandations.

Cadre Théorique

A. Présentation des articles de référence

MONTMARQUETTE, 2020

(Montmarquette, 2020) présente une revue de littérature très détaillée sur les déterminants du choix du domaine d'études universitaires. Il a d'abord dégagé les deux principaux cadres conceptuels dans lesquels s'inscrivent habituellement les différentes approches : l'un qui relève de la psychologie et l'autre du modèle économique du capital humain. Si le premier fait référence au type de personnalité ((réaliste, expérimental, artistique, social, entreprenant et conventionnel), l'autre stipule que : « *L'idée standard est que les étudiants choisiront un domaine spécifique si la valeur actuelle attendue de l'utilité à vie pour choisir ce domaine d'études est supérieure à la valeur attendue d'un autre domaine* ».

Il en ressort de son travail des conclusions assez pertinentes permettant d'avoir une vue globale sur l'ensemble des déterminants du choix du domaine d'études universitaires telles que :

- la valeur attendue de consommation de l'éducation joue un rôle majeur pour prédire la probabilité déclarée par les étudiants de poursuivre leurs études à plein temps ;
- les individus sont moins influencés par les niveaux de revenus initiaux dans les professions liées aux différents domaines d'études, mais davantage par le flux de revenus que ces professions devraient générer ;
- la probabilité de réussite perçue dans un domaine d'études influe sur le domaine choisi ;
- la présence d'un compromis entre le risque d'échec et le rendement économique de l'éducation ;
- l'avantage comparatif de réussite influe sur le choix observé de la concentration d'études ;
- le lieu de résidence a un impact sur le choix des domaines d'études ;

MONTMARQUETTE, BELLEMARE & KESER, 2000

(Montmarquette, Bellemare, & Keser, 2000) ont fait ressortir les considérations financières et institutionnelles de la décision d'études au Québec. Il a été observé que :

- a. les étudiants choisissent leur filière d'études en fonction de leur souci de maximiser leur bien-être qui est une fonction directe de leur espérance de revenu sur un horizon de temps défini ;
- b. le flux de revenus futurs anticipés par formation est plus déterminant du choix de filières que les revenus à toucher immédiatement après la graduation ;
- c. une étude américaine ayant mis de côté toute considération de revenus a conclu que les préférences des étudiants, leurs anticipations de revenus et la probabilité de trouver un emploi après les études seraient les facteurs résiduels les plus importants à analyser.

En conclusion de ce papier, le choix des filières est fonction de certains facteurs tels :

- les revenus espérés après la graduation et la projection des revenus futurs anticipés (ces variables captent les écarts de revenus entre les diplômés ainsi que le risque de compléter avec succès ou non un programme)
- certaines préférences comme le nombre d'enfants anticipés, la participation sur le marché du travail, désir de mobilité

Magali JAOUL-GRAMMARE

Magali JAOUL-GRAMMARE, dans « Stratégie des étudiants, choix d'orientation et problèmes d'engorgement : Une analyse quasi-expérimentale », explique les facteurs qui peuvent influencer sur le choix de l'université à partir d'un protocole expérimental fait sur différents sujets. Vu que les étudiants ne font pas de calculs explicites du taux de rendement mais ils ont tendance à réagir aux changements et aux évolutions affectant ainsi les diverses filières de formation. De ce fait il se base sur l'hypothèse de Diebolt « théorie de l'engorgement » (2001) qui stipule que le choix des étudiants dépend des avantages comparatifs en termes de salaires espérés et de travail disponible.

L'expérimentation se fait à partir de trois jeux avec des gains et un nombre de joueurs variables ; et qui sera suivi d'un questionnaire verbal. Les sujets expérimentés choisis, font face à trois contraintes : un nombre de places limité, un gain variable et une dotation variable. Ces différentes contraintes vont servir à étudier le comportement des individus face au nombre de places disponibles sur le marché du travail face à l'évolution des salaires d'une part, et d'autre part, à tenir compte de leur origine sociale.

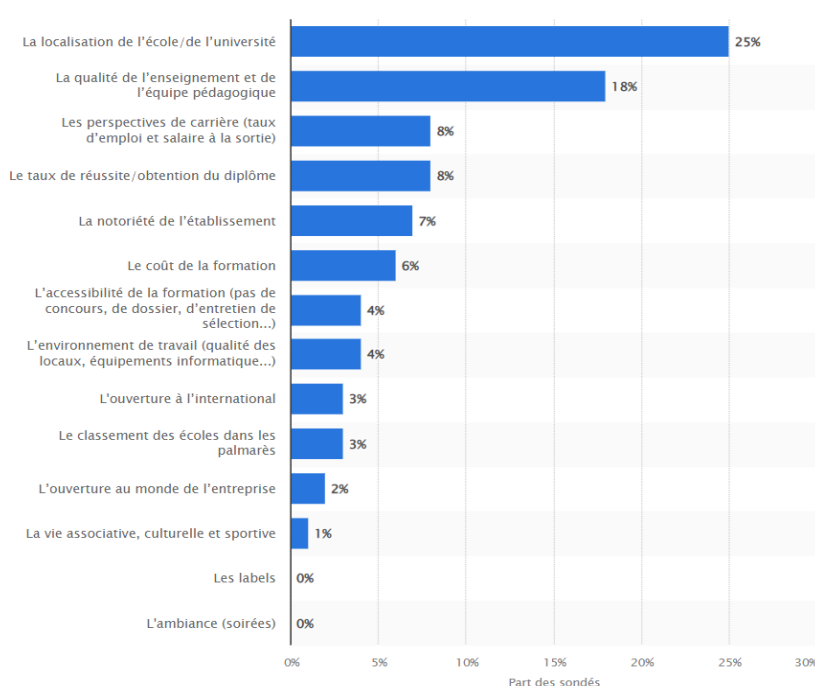
Les résultats de l'expérience comparé aux prévisions du jeu révèlent que ces derniers tiennent compte du nombre de places et du gain dans leur choix. Toutefois, ils n'anticipent point le comportement des autres afin d'adapter leur choix. Ce qui semble confirmer l'hypothèse du

« modèle théorique de l'engorgement » même si le questionnaire verbal nuance dans une certaine mesure les résultats.

Statista Research Department

En 2021, Statista Research Department a publié les statistiques indiquant l'avis des jeunes français concernant les critères les plus importants pour le choix de l'établissement de formation en avril 2018. Ces données présentées dans le tableau ci-contre nous enseignent sur la haute importance de la localisation, la qualité de l'enseignement et les perspectives de carrière dans le choix du domaine d'études des jeunes français.

Graphe 1. Les critères les plus importants au moment de choisir l'établissement de formation



Statista Research Department, 2021

De moindre importance, d'autres facteurs sont également pris en compte par les étudiants dans le choix de leur domaine d'études comme le taux de réussite dans la formation, la notoriété de l'établissement ou encore son ouverture à l'international.

B. Hypothèses

A partir de ses articles, nous avons décidé de baser notre travail sur les trois hypothèses suivantes :

- H1 : le choix des étudiants dépend des avantages comparatifs en termes d'emplois disponibles sur le marché

- H2 : le choix des étudiants dépend des avantages comparatifs en termes de salaires espérés
- H3 : le choix de l'université des étudiants dépend de la notoriété de l'université

Analyse empirique

A. Analyse descriptive

Présentation des données

A partir des hypothèses énoncées précédemment, nous avons pu trouver des variables pertinentes pour notre étude. En effet de ces hypothèses et ces articles, nous avons noté des critères de choix d'université qui revenaient le plus souvent. Ces critères de choix peuvent ainsi être assimilés à des variables de décision. Négliger ces variables reviendrait alors à biaiser notre étude et sans elles nous ne pouvons que nous attendre qu'à des faux résultats.

Notre travail devra se faire alors sur une base de données regroupant la plupart de ces variables. Ainsi, avons-nous opté pour une base de données trouvée sur le site de data-gouv. Il s'agit d'une étude portant sur l'évolution des étudiants ayant obtenu un master au sein de diverses universités de France. En effet à la fin de chaque année les universités mènent une enquête auprès de leurs diplômés pour connaître leur devenir professionnel, le niveau d'insertion professionnelle des étudiants ayant un suivi un parcours et leur temps d'attente avant un tout premier emploi. De plus à cela on retrouve alors des informations sur les différents parcours des universités. Cette base très complète, nous permettra de vérifier nos hypothèses dans le sens où elle renferme les variables adéquates pour notre étude.

Quelles sont alors ces variables ?

Pour les besoins de notre étude, nous avons retenu dans la base de données, les variables suivantes :

Variable expliquée :

Modalités	Universités	Position classement Shanghai
U0	Université Paris-Saclay	1
	Sorbonne Université	2
U1	Université de Paris	4
	Aix-Marseille	5-7
U2	Grenoble Alpes	5-7
	Strasbourg	5-7

Modalités	Universités	Position classement Shanghai
U3	Montpellier	8
	Lyon 1 - Claude Bernard	9
U4	Bordeaux	9-11
	Toulouse 3 - Paul Sabatier	12-15

Etablissement : Vu que l'on s'intéresse *aux déterminants des choix d'université par les étudiants*, nous avons choisi cette variable comme variable explicative. Elle comporte la liste des universités publiques de France. Ne pouvant travailler sur la totalité des universités de France (plus de 60 universités disponible), nous avons décidé à partir du classement Shanghai, de sélectionner les 10 meilleures universités de France et les regrouper par binôme. Après un travail de codification, nous nous retrouvons avec la variable établissement renommé code_etab qui prend 5 modalités. On retrouve les différentes modalités dans le tableau ci-dessous.

Variables explicatives

Comme dit précédemment , pour le choix de nos variables nous nous sommes basés sur les critères de décision dans le choix de l'université qui sont ressortis de nos articles et hypothèses. Nous sommes partis ainsi de l'appréhension que le choix de l'université dépendait de 7 variables .Ces variables sont représentées dans le tableau ci-dessous :

Variables	Type de variable	Signification	Exemples de modalités
Code du domaine	Qualitative Multinomial	Parcours , formation de chaque universités	0: DEG 1: LLA 2: MEEF 3: SHS 4: STS
Situation	Qualitative Binomial	Temps d'attente avant le premier emploi	0 : 18 mois 1 : 30 mois
Poids de la discipline	Quantitative	Valeur de chaque formation sur le marché de l'emploi	40
Taux d'insertion	Quantitative	Taux d'étudiants ayant réussi une insertion	90%

Variables	Type de variable	Signification	Exemples de modalités
		professionnelle par parcours et université	
Emplois cadre ou professions intermédiaires	Quantitative	Nombre d'étudiants ayant un emploi cadre/intermédiaires par parcours et université	542
Emplois stables	Quantitative	Nombre d'étudiants ayant un emploi stable par parcours et université	56
Salaire brut annuel estimé	Quantitative	Salaire annuel estimé des étudiants ayant suivis une formation	20 000 Euros

Dans l'ensemble, nous avons des résultats très pertinents pour la statistique descriptive. En effet, nous notons pour la moyenne des variables des valeurs très élevées, notamment 73.95% pour les emplois stables et 94.02% pour les emplois à temps plein. De plus le salaire brut annuel estimé prend des valeurs relativement élevées et varie entre 23 400 et 43 100 euros pour une médiane de 31 900 euros. Au vu de ces premiers résultats, cela semble rassurant pour un nouvel étudiant. En effet, ce dernier n'aura pas de souci d'insertion professionnelle dans la plupart des universités de notre échantillon. Ces résultats peuvent tout aussi conforter l'idée que le milieu scolaire intervient de manière significative dans le milieu professionnel.

Si nous poursuivons notre analyse avec le skewness, nous notons pour le taux d'insertion une valeur négative (- 0.436260) ce qui signifie que la distribution est décalée vers la droite avec sa queue sur le côté gauche. C'est une distribution biaisée à gauche car la moyenne de la variable est aussi inférieure à sa médiane. Le skewness du poids de la discipline est quant à lui positif (1.871810) ce qui signifie que la distribution est biaisée à droite.

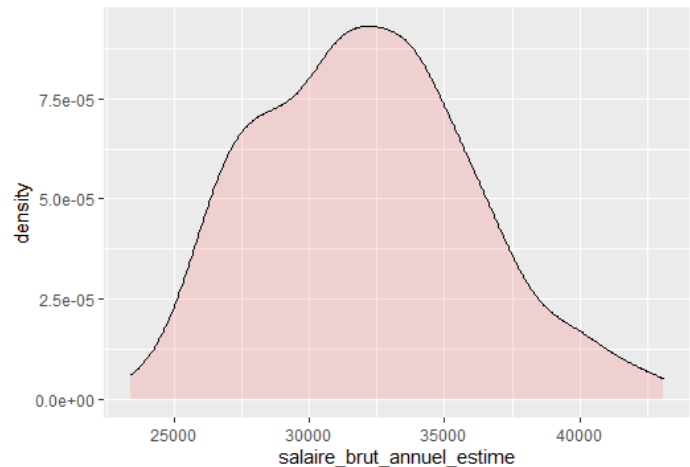
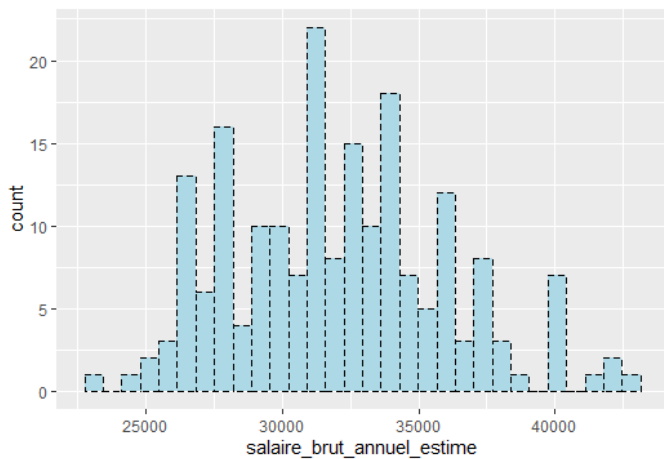
Le kurtosis de la variable emplois cadre est de 3.396747 ce qui signifie que sa distribution est trop élevée contrairement à celui de la variable emplois stables qui est négatif (0.715562) qui signifie que sa distribution est trop plate.

L'écart type pour la variable poids de la discipline est proche de sa moyenne, ce qui n'est pas le cas pour le reste des variables pour lesquelles l'écart type est très éloigné de la moyenne signifiant que les données sont réparties sur une large plage de valeurs.

	Poids dicip.	taux d'insert.	emplois cadre	emplois stables	Emplois temps p.	Slre brut ann.estim.
<i>Minimum</i>	2.000000	75.000000	55.000000	41.000000	62.000000	23400
<i>Maximum</i>	100.000000	100.000000	100.000000	100.000000	100.000000	43100
<i>1. Quartile</i>	7.000000	87.000000	89.000000	64.000000	92.000000	28975
<i>3. Quartile</i>	28.000000	94.000000	98.000000	85.000000	98.000000	34400
<i>Mean</i>	24.132653	90.433673	92.107143	73.959184	94.025510	32084
<i>Median</i>	13.000000	91.000000	94.000000	74.500000	96.000000	31900
<i>Sum</i>	4730.000000	17725.000000	18053.000000	14496.000000	18429.000000	6.288500e+06
<i>SE Mean</i>	1.918803	0.370630	0.566341	1.008022	0.484152	2.811718e+02
<i>LCL Mean</i>	20.348382	89.702716	90.990203	71.971158	93.070665	3.152966e+04
<i>UCL Mean</i>	27.916924	91.164631	93.224082	75.947209	94.980356	3.263871e+04
<i>Variance</i>	721.633595	26.923783	62.865385	199.157300	45.942936	1.549529e+07
<i>Stdev</i>	26.863239	5.188813	7.928769	14.112310	6.778122	3.936405e+03
<i>Skewness</i>	1.871810	-0.436260	-1.621692	-0.183993	-2.143541	3.118680e-01
<i>Kurtosis</i>	2.488044	0.085939	3.396747	-0.715562	5.571682	-3.172370e-01

Histogramme et fonction de densité :

- Salaire brut annuel estimé

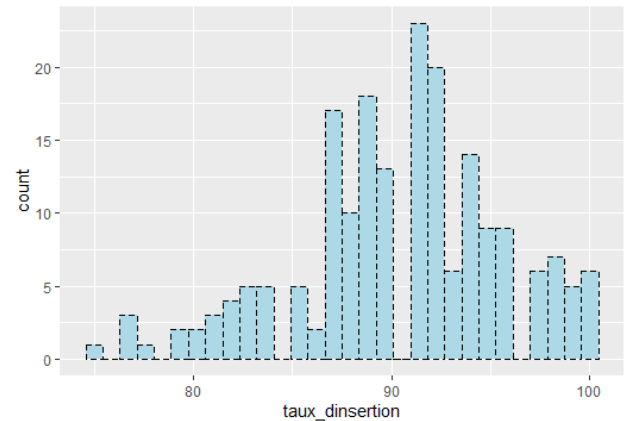
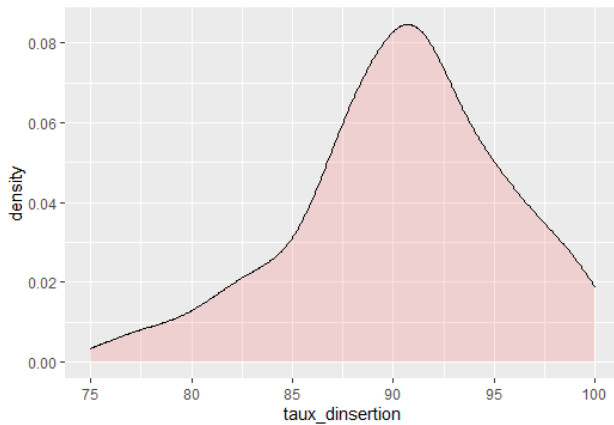


Interprétation :

Nous notons une répartition assez hétérogène du salaire. En effet, le montant du salaire annuel estimé des étudiants varie entre 25 000 euros et 40 000 euros par an. De plus, nous remarquons un pic autour des montants 30 000 à 35 000. Par contre on retrouve peu d'effectif au-delà des 40 000 euros par an. A priori les études dans les meilleures universités peuvent sous-entendre voir un salaire assez élevé. Nous ne pouvons pas hélas confirmer cela sans des tests économétriques, mais l'histogramme a tendance à nous amener cette affirmation. De plus, il est important de préciser que ces effectifs regroupent toutes les universités et toutes les formations sans distinction. Il est alors normal que des universités proposant des formations avec un salaire en fin d'études assez élevé enregistrent de grands effectifs contrairement à d'autres.

Il serait alors intéressant dans le reste de notre étude de voir comment cette variable évolue en fonction de l'université choisie.

- Taux d'insertion professionnelle

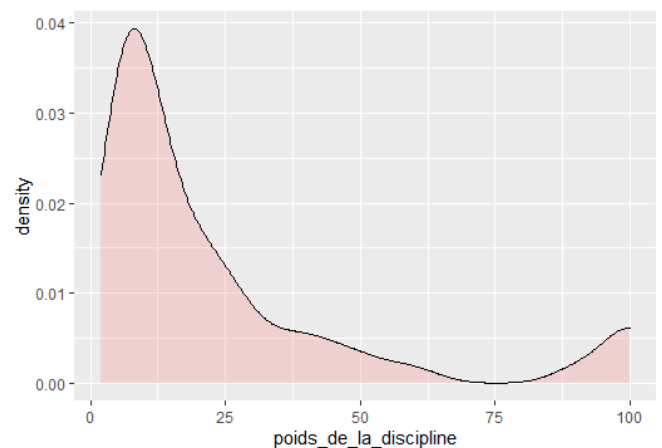
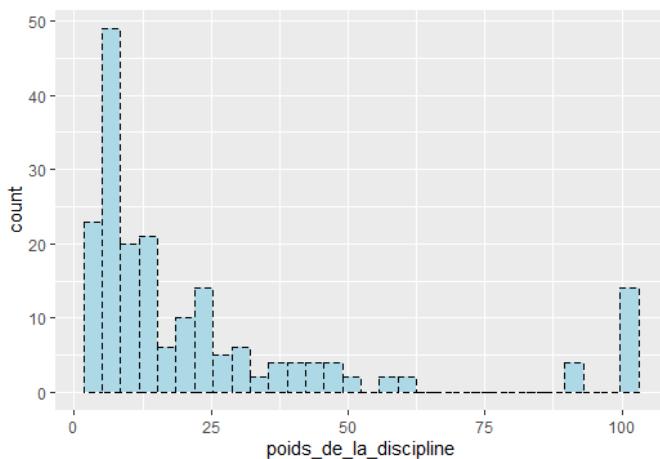


Interprétation :

Un premier regard sur ces graphiques nous permet de noter que le taux d'insertion, pour la plupart des universités est relativement élevé. Cela peut tout aussi être retrouvé dans la statistique descriptive où nous retrouvons une valeur minimale de 75%. Si nous entrons en détail, la plupart des universités enregistrent des taux d'insertion aux alentours 85-95%. De plus, même au-delà de ce pic d'effectif, le taux d'insertion reste plus élevé. De ce fait, nous notons qu'une certaine opportunité professionnelle garantie peut découler de choisir les meilleures universités.

Cette variable est confrontée aux mêmes problèmes que précédemment d'où l'impossibilité de tirer des conclusions à priori.

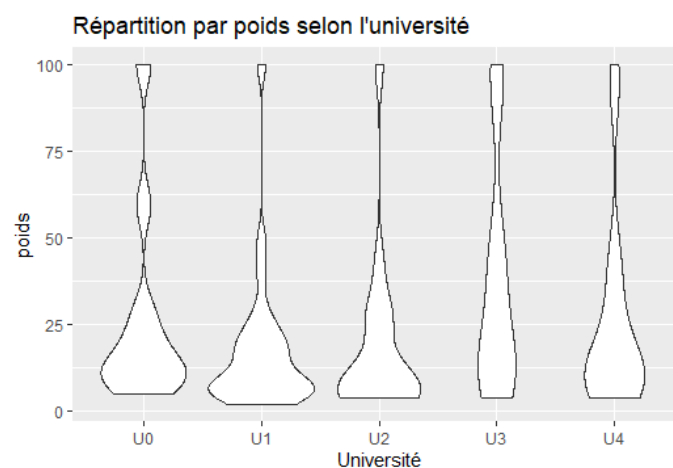
- Poids de la discipline :



Interprétation :

Contrairement aux précédentes variables, le poids de la discipline enregistre un plus grand effectif vers les valeurs les plus faibles. Ce résultat est assez intéressant dans le sens où nous avons choisi les meilleures universités de France, il aurait été normal de voir des valeurs assez grandes. Une raison à cela peut être du fait que même si les universités sont les meilleurs dans un classement général, elles ne sont pas forcément le meilleur dans des certaines spécialités pris individuellement. Pour exemple l'université d'Orléans fortement coté pour sa formation Droit éco gestion : Statistique ne se trouve pas dans le top 20 dans le classement Shangai et donc à échapper à notre sélection.

- Etablissement



Interprétation :

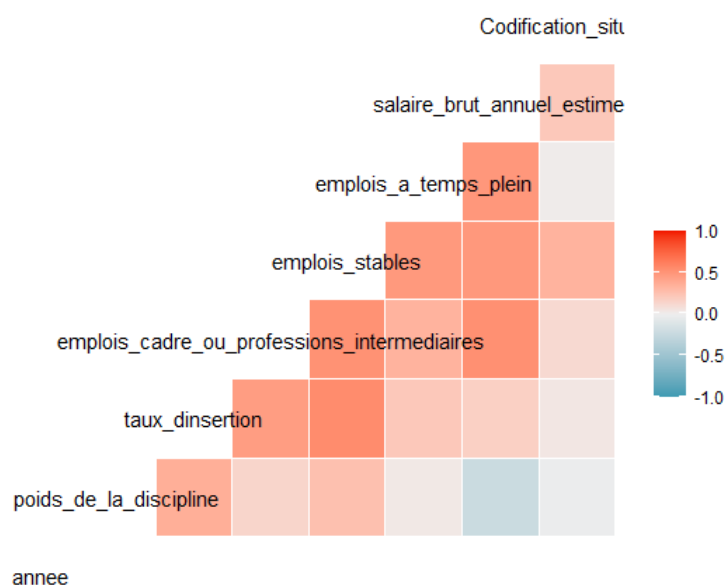
➤ Répartition par poids :

Comme l'histogramme de poids de la discipline l'a montré, les meilleures universités que nous avons choisies enregistrent des poids de formation relativement faibles. Les universités du groupe 3 et 4 apparaissent avoir de meilleurs poids de formation que celles des groupes 1 et 2. De plus la totalité des groupes présentent de grands effectifs pour les poids de formation faible. En effet, au fur et à mesure que le poids de formation diminue l'effectif à tendance à augmenter pour la totalité des universités.

➤ Répartition par taux d'insertion :

Contrairement aux résultats précédents, les meilleures universités ont tendance à avoir des taux d'insertion très élevés. En effet nous pouvons que pour l'ensemble des universités les plus grands effectifs tournent autour des taux d'insertion de 85-95%. Notons juste que les 3 premiers groupes, à savoir le groupe des meilleures universités à tendance à enregistrer des taux d'insertion assez faible en dessous de 80% , ce qui n'est pas le cas pour les deux derniers groupes. De plus, le tout dernier groupe enregistre un effectif de taux d'insertion plus élevé que le premier groupe.

Matrice de corrélation



Interprétation :

La matrice de corrélation indique les valeurs de corrélation. Elles mesurent le degré de relation linéaire entre chaque paire de variables. Les valeurs de corrélation sont comprises entre -1 et +1. Si les deux variables ont tendance à augmenter et à diminuer en même temps, la valeur de corrélation est positive. Dans notre situation, nous pouvons noter une corrélation positive entre la plupart des variables. Cependant il y'a une corrélation légèrement négative qui ressort entre la variable poids de la discipline et emplois à temps plein. Au-delà de cela nous enregistrons une corrélation plus ou moins nulle entre les variables emplois temps plein et salaire brut annuel estimé, les variables poids de la discipline et taux d'insertion avec la variable situation (30 ou 18 mois). Mis à part ces derniers, nous pouvons voir une corrélation positive pour les autres variables tournant autour de 0.5

B. Présentation du modèle :

Le modèle et choix de la variable de référence

Afin d'expliquer les déterminants du choix d'université par les étudiants, nous avons opté pour un modèle de régression logistique multinomiale ordonnée. Suivant les variables explicatives considérées, notre modèle peut s'écrire :

$$\left\{ \begin{array}{l} y_i = U_0 \text{ si } y_i^* \leq C_1 \\ y_i = U_1 \text{ si } C_1 < y_i^* \leq C_2 \\ y_i = U_2 \text{ si } C_2 < y_i^* \leq C_3 \\ y_i = U_3 \text{ si } C_3 < y_i^* \leq C_4 \\ y_i = U_4 \text{ si } y_i^* > C_4 \end{array} \right.$$

Avec :

$$y_i^* = \beta_0 + \beta_1 \text{ poids_de_la_discipline} + \beta_2 \text{ taux_dinsertion} + \beta_3 \text{ emplois_a_temps_plein} + \beta_4 \text{ emplois_stables} + \beta_5 \text{ emplois_cadre_ou_professions_intermediaires} + \beta_6 \text{ Codification_situation} + \beta_7 \text{ code_domaine} + \varepsilon_t$$

Nous également fait choix de U_2 comme catégorie de référence étant la catégorie centrale de notre classement.

Le modèle de régression multinomiale ordonnée obtenu à partir de la librairie « nnet » nous a permis d'avoir le tableau ci-après rassemblant ainsi les coefficients et leurs écarts-type entre parenthèses.

(multinom(formula = etabRef ~ poids_de_la_discipline + emplois_a_temps_plein + emplois_cadre_ou_professions_intermediaires + code_domaine, data = data))

```
> stargazer(modele, type="text", out="multi1.htm")
```

	Dependent variable:			
	U0 (1)	U1 (2)	U3 (3)	U4 (4)
poids_de_la_discipline	0.037* (0.021)	-0.002 (0.022)	0.040* (0.021)	0.033 (0.020)
taux_dinsertion	-0.211*** (0.063)	-0.171*** (0.058)	-0.132* (0.072)	-0.097 (0.063)
emplois_a_temps_plein	-0.032 (0.052)	-0.206*** (0.052)	-0.058 (0.064)	-0.024 (0.061)
emplois_stables	0.082*** (0.030)	0.067** (0.029)	0.060* (0.031)	0.049* (0.029)
emplois_cadre_ou_professions_intermediaires	0.169*** (0.058)	0.029 (0.044)	-0.072 (0.060)	0.055 (0.053)
Codification_situation	-0.960 (0.584)	-0.467 (0.564)	-0.242 (0.618)	-0.430 (0.578)
code_domaine1	2.326* (1.401)	-2.440* (1.278)	-30.568*** (0.000)	-22.108*** (0.000)
code_domaine2	-2.598 (2.062)	-1.367 (2.173)	-1.862 (2.081)	-3.070 (2.079)
code_domaine3	1.799* (0.955)	-1.947** (0.944)	-14.590*** (0.00001)	0.314 (0.978)
code_domaine4	-0.050 (0.755)	-0.788 (0.693)	1.385 (0.878)	0.318 (0.719)
Constant	0.117 (5.148)	28.924*** (5.128)	18.504*** (0.810)	2.171 (5.665)
Akaike Inf. Crit.	613.714	613.714	613.714	613.714
Note:	*p<0.1; **p<0.05; ***p<0.01			

Tableau : modèle logit multinomial sans itérations des paramètres

Il en ressort que plus le taux d'insertion est significatif pour les catégories U_0 et U_1 , significatif à 10% pour la catégorie U_3 et non significatif pour la catégorie U_4 . Les signes négatifs de ce paramètre nous permettent de conclure que plus le taux d'insertion est élevé moins les étudiants ont la probabilité de choisir une catégorie autre que U_2 . Cela peut s'expliquer par le fait qu'à opportunité égale l'étudiant priorise le taux d'insertion plutôt que la catégorie. En effet, l'étudiant peut estimer que U_2 est plus prestigieux que U_3 et U_4 tout en étant moins contraignant que les plus prestigieux U_0 et U_1 .

Par ailleurs, plus l'assurance d'avoir un emploi stable à la fin de son cursus est élevée, plus la probabilité de choisir les universités des catégories U_0 et U_1 plutôt que U_2 augmente. Il en va de même pour l'emploi cadre ou professions intermédiaires pour lesquels la probabilité de choisir les universités de la catégorie U_0 augmente par rapport à la catégorie U_2 .

D'autre part, en ce qui a trait aux domaines d'études, la probabilité de choisir la catégorie U_0 augmente par rapport à U_2 tandis que les probabilités de choisir U_1 et U_3 diminuent par rapport à U_2 pour les étudiants qui souhaitent postuler en sciences humaines et sociales (code_domaine3),

Toutefois, notons que le temps d'embauche après l'obtention du diplôme n'est nullement significatif dans le choix de catégorie d'université.

Matrice de confusion

	U2	U0	U1	U3	U4	Total
U2	11	1	6	1	3	22
U0	9	25	11	6	8	59
U1	4	11	24	5	10	54
U3	2	2	8	15	8	35
U4	6	5	3	4	8	26
Total	32	44	52	31	37	196

Le modèle de régression logistique multinomiale spécifié présente un taux global de précision de 42%¹. Ce qui implique que les faux positifs et les faux négatifs sont très élevés dans le modèle.

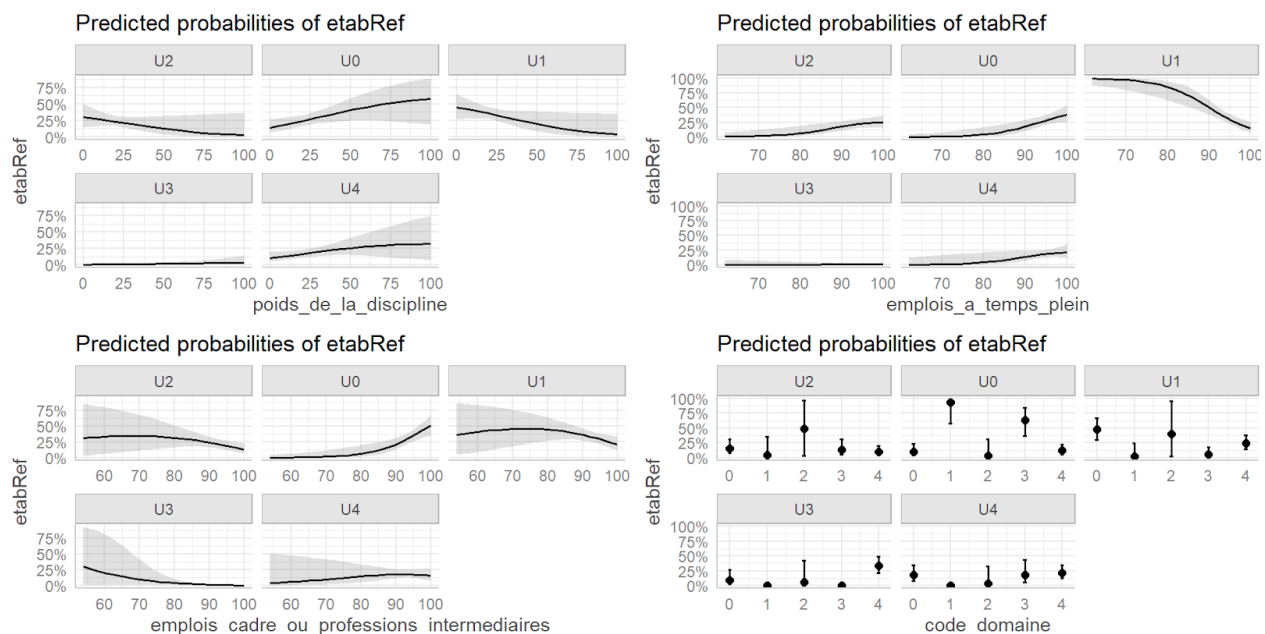
Effets marginaux

Les effets marginaux nous permettent de voir plus en détail ce que le résultat de régression fournit de plus. En effet, ils nous renseignent sur l'impact que pourrait avoir les variables explicatives sur le choix du domaine.

Notons bien qu'il est plus intéressant d'analyser les effets marginaux sur la probabilité prédite plutôt que sur les coefficients ou les odds-ratios.

Les graphes ci-dessous représentent les effets marginaux des différentes variables pour chaque modalité. La lecture graphe par graphe nous permet d'étudier l'impact d'une variable explicative sur chaque modalité.

¹ Le taux de précision est obtenu en divisant le nombre de bonnes réponses (la somme des valeurs se trouvant sur la diagonale) au nombre total d'observations (196)



La catégorie U2 étant toujours la modalité de référence, on constate que le poids de la discipline augmente la probabilité de passer de U2 à U0 et diminue pour passer de U2 à U1. Concernant l'emploi à temps plein, elle diminue fortement la probabilité de U2 à U1. Mais augmente pour passer de U2 à U0.

Conclusion

Eu égard à tout ce qui précède, notre analyse cherchait à mettre en relief l'impact de différentes variables dans le choix de l'université par les étudiants. En effet, notre intérêt s'est tourné vers cette problématique en raison de notre situation d'étudiants. De plus, il est intéressant de noter comment, un choix qui peut conditionner toute une vie peut se faire. Cette étude que nous avons menée peut-être poussé vers une analyse psychologique du choix d'individus. Pour se faire nous sommes partis d'un modèle multinomial ordonné. Nous avons de ce fait, modéliser, le choix de diverses universités en fonction de plusieurs variables notamment taux d'insertion, formation, poids de la formation etc. En effet, une étape principale, était de choisir les variables adéquates pour comprendre ce choix d'université. Nous nous sommes ainsi reposés sur divers articles scientifiques. Ces dernières ont su nous menés sur la voie à suivre, il a été facile alors d'en tirer des hypothèses sur lesquelles basées notre étude.

Pour comprendre la portée de nos variables, nous avons réalisé en premier lieu une statistique descriptive. Les résultats issus de cette première approche ont été au-delà de nos attentes. En effet nous avons noté qu'au-delà du classement des universités, la meilleure université du classement n'enregistrait pas forcément les meilleures valeurs au niveau des variables explicatives. Pour exemple, avec la variable taux d'insertion professionnelle, la meilleure université n'est pas celle qui enregistre de meilleur chiffre. Ce phénomène de différence entre le classement et les résultats des variables explicatives se retrouve au niveau de la plupart de nos variables explicatives. En effet nous avons à travers cela pu envisager la possibilité que malgré la position dans le classement, les universités du top ne sont pas forcément les meilleurs dans tous les domaines. Il serait intéressant de ce fait de mener une étude annexe sur la réalisation du classement auquel nous avons eu recours.

Par la suite, nous nous sommes attaqués l'essentiel du problème, à savoir notre estimation. Nous avons noté encore une fois de plus des résultats très hétérogènes. Tout d'abord la significativité des variables variait d'une variable endogène à une autre. Comme exemple, on retrouve le taux d'insertion très significatif pour la modalité U0 et U1, mais par contre elle l'est moins, voire pas du tout pour les deux dernières modalités. Nous avons pu conclure que l'importance des différentes variables variait drastiquement d'une université à une autre. En effet un étudiant se tournant vers les universités de la modalité U0 prendra en compte essentiellement le taux d'insertion, le nombre de diplômés ayant un emploi stable, alors qu'un autre tourné vers les universités de la modalité U4 non, celui-ci s'intéressera beaucoup plus à

la formation. Il en résulte que les choix des étudiants ne s'expliquent pas forcément que par ces paramètres (42% de précision). D'autres considérations plus personnelles ou liées à des paramètres non intégrés au modèle peuvent intervenir dans le choix de leur domaine d'études et/ou de l'université.

Outre les résultats de notre étude, il serait intéressant d'inclure comme du précédemment l'importation des considérations personnelles. En effet la notion, de vision et d'objectif peuvent varier d'un individu à un autre, et en fonction de ces dernières le choix de l'université peut tout aussi être fait. Nous voulons alors montrer qu'il faudrait tenir compte du fait qu'un étudiant, en général décide d'embrasser les formations qui conviennent à sa conviction et à ses perspectives. Par ailleurs il pourrait y avoir un impact parental qui pourrait induire le choix des futurs étudiants. Ainsi, autant de variables qui peuvent impacter ce choix de formation. La complexité dans cette tâche de modélisation du choix d'université se pose alors. Pour une suite il serait intéressant de voir comment aborder problème et ainsi intégrer dans l'estimation l'essentiel des variables endogènes pouvant avoir un impact sur le choix de l'université. En somme, notre étude n'a fait qu'une ébauche du travail qu'il reste à faire. Nous avons pu mettre en exergue le problème de manquement de variables adéquats mais aussi celui d'hétérogénéité entre classement et différentes valeurs. Il serait intéressant d'en tenir compte dans une seconde approche qui se voudrait beaucoup plus complète

Bibliographie

- Montmarquette, C. (2020). *Les déterminants du choix du domaine d'études universitaires : Une revue de la littérature et identification de pistes d'interventions*. CIRANO et Université de Montréal.
- Montmarquette, C., Bellemare, C., & Keser, C. (2000). *Le Choix des Filières d'Études au Québec : Situation Actuelle et Revue de la Littérature*. Montréal: CIRANO.
- Statista. 'Critères de choix d'établissement d'études selon les jeunes français 2018'. Accessed 18 March 2022.
<https://fr.statista.com/statistiques/619063/choix-etudes-jeunes-criteres-les-plus-importants-france/?fbclid=IwAR0a0qvHIEZK2K1uqXguUSwhmM26SsCc6NMyi6Fk6XJorRPAig57TZtFyME>.
- Jaoul, Magali. 'Higher education and labour markets: an econometric analysis of the "glutting" theory'. *Economie prevision* 166, no. 5 (2004): 39–57.c'Sr500.Pdf'. Accessed 18 March 2022.
https://www.newyorkfed.org/medialibrary/media/research/staff_reports/sr500.pdf.
- WISWALL, MATTHEW, and BASIT ZAFAR. 'Determinants of College Major Choice: Identification Using an Information Experiment'. *The Review of Economic Studies* 82, no. 2 (291) (2015): 791–824.

Table des matières

Sommaire	0
Introduction.....	1
Cadre Théorique.....	3
A. Présentation des articles de référence.....	3
B. Hypothèses	5
Analyse empirique	7
A. Analyse descriptive	7
Présentation des données	7
B. Présentation du modèle :	15
Conclusion	19
Bibliographie.....	21
Table des matières.....	22

Annexe 1. Packages R utilises

PACKAGES	COMMANDE	SYNTAXE	FINALITES
fbasics	basicStats	basicStats(x)	Dégager les propriétés statistiques
readxl	read_excel	read_excel("repertoire",col_types =c())	Importer une base de donnée excell
ggplot2	ggplot	ggplot(data, aes(x=variable))	Permet de faire un plot sous différentes formes
~	geom_histogram	geom_histogram(color= , fill= ,linetype=)	Dégager un histogramme
~	geom_density	geom_density(alpha= , fill=)	Dégager une fonction de densité
~	ggcorr	ggcorr(data)	Dégager une matrice de corrélation
nnet			Librairy pour faire une estimation multinomiale
	multinom	multinom(y~x+x,data=base)	Faire une estimation multinomiale
stargazer		stargazer(modele, type="text", out="multi1.htm")	Significativité des variables
effects	marginal_effects	marginal_effects(modele)	Dégager les effets marginaux
	ggcoef_multinom	ggcoef_multinom(modele2,exponentiate = TRUE)	Générer une matrice de confusion
effects	allEffects()	allEffects(modele)	Dégager les effets marginaux