

Université d'Orléans

## QUE SONT DEVENUS LES SORTANTS DU SYSTEME EDUCATIF DE LA REGION CENTRE ?

Problématique :

*Identification des caractéristiques influençant la probabilité  
d'avoir un contrat à durée indéterminée lors du premier emploi.*

MASTER I ESA



Université d'Orléans

Econométrie et Statistiques Appliquées

Auteurs : CHEVREAU Antoine  
GODIN Sylvain

Directeur de mémoire : HURLIN Christophe

## Sommaire :

### Introduction.

#### I Statistiques descriptives.

1. Influence du sexe.
2. Influence du statut marital.
3. Influence de la nationalité.
4. Influence de la taille de l'entreprise.
5. Influence du type d'intermédiaire.
6. Influence du secteur d'activité.

#### II Influences de différentes variables explicatives.

- 1) Modèle 1.
- 2) Modèle 2.
- 3) Modèle 3.
- 4) Modèle 4.
- 5) Modèle 5.

#### III Choix du modèle optimal.

### Conclusion.

### Annexes

## Introduction :

Les données sur lesquelles nous allons travailler proviennent d'une enquête menée par le CEREQ (Centre d'Etudes et de Recherches sur les Qualifications) dont l'intitulé est « Que sont devenus les sortants du système éducatif de la région Centre ? ». Notre base de données comprend 2197 individus qui ont répondu à différentes questions réparties en 8 grandes catégories :

- 1 :« Votre état civil »
- 2 :« Votre formation »
- 3 :« Votre situation actuelle »
- 4 :« Aujourd'hui vous avez un emploi »
- 5 :« Aujourd'hui, vous êtes sans emploi »
- 6 :« De la fin de vos études à votre situation actuelle »
- 7 :« Description du premier emploi occupé »
- 8 :« Les différentes situations que vous avez rencontrées ».

A l'aide de cette base, nous avons choisi de déterminer la probabilité d'obtenir un CDI dès son premier emploi c'est-à-dire que nous allons nous occuper uniquement des catégories 1, 2, 7. On omet le cadre 5 car seules les personnes ayant trouvé un contrat dans les entreprises nous intéressent, les cadres 3 et 6 car ceux sont des questions intermédiaires visant à sélectionner les personnes ayant un emploi ou non et les personnes ayant occupé un emploi ou non par le passé.

Notre objectif est de déterminer la probabilité d'avoir un CDI comme premier emploi à la fin des études en fonction :

- Des caractéristiques de l'individu situées dans les cadres 1 et 2 (« état civil » et « formation »).
- Des caractéristiques de son premier emploi occupé se trouvant dans le cadre 7.

Notre travail consistera dans un premier temps à essayer de sélectionner les individus qui vont nous intéresser pour notre problématique. Puis dans un second temps, il nous faudra sélectionner les variables qui nous paraissent les plus pertinentes parmi celles présentées dans les deux catégories ci-dessus. Enfin, nous allons essayer de construire un modèle qui prédit le mieux les influences des différentes variables sur la probabilité d'avoir un CDI comme premier emploi.

Dans une première partie, nous allons voir quelques statistiques descriptives pour nous familiariser avec notre base de données.

Puis dans une seconde partie, nous allons construire et interpréter différents modèles afin de trouver un modèle pertinent en effectuant une analyse économétrique.

## **I. Statistiques descriptives.**

### **Statistiques descriptives générales : voir annexe 1.**

Via les statistiques descriptives du tableau 1 on peut dire que notre base est constituée de 61,81% d'hommes et donc de 38,19% de femmes donc les individus ayant répondu au questionnaire sont majoritairement des hommes. Par ailleurs, notre base est constituée de personnes ayant la nationalité française à 98,45%, les personnes à nationalité étrangères sont donc en minorité dans notre échantillon. D'autre part, les personnes interrogées sont célibataires, divorcé(e), veuf(ve) à hauteur de 73,19% et vivent maritalement ou en concubinage à 26,81%. Au moment de l'enquête, 75,33% déclarent avoir un emploi contre 24,67% qui se déclarent au chômage.

Passons aux statistiques descriptives concernant le premier emploi occupé. En ce qui concerne le type de contrat signé pour le premier emploi, 31,32% des personnes ont eu un CDI comme premier emploi et 37,78% ont eu un CDD (pour ne citer que les deux formes principales de contrat).

Par ailleurs, 27,51% ont été employés dans une entreprise ayant moins de 10 salariés, 28,57% ont trouvé un travail dans une entreprise ayant 10 à 49 salariés et 17,99% ont été embauchés dans une grande entreprise de 200 salariés et plus. On a donc environ 55% des personnes interrogées qui ont trouvé un premier emploi dans une PME (petite et moyenne entreprise).

Les principaux secteurs d'activités dans lesquels les personnes ont trouvé un premier emploi sont l'industrie mécanique, métallurgique pour 11,48% des personnes, autres industries (bois, textile, chimie, graphique,...) pour 11,90% et Hôtellerie restauration pour 9,88% des cas.

Quant à la nature de l'entreprise, 84,59% des premiers emplois ont été trouvés dans des entreprises privées. On peut noter que plus de 66% des premiers emplois ont été trouvés grâce à des démarches personnelles et que ceux-ci sont à plus de 80% des emplois à temps plein.

Pour observer l'impact des différentes variables sur la probabilité d'avoir un CDI comme premier emploi, on a tout d'abord créé une variable dichotomique appelée cdi telle que :

$$\text{cdi}_i = \begin{cases} 1 & \text{si la variable } g2 = '1' \\ 0 & \text{sinon} \end{cases}$$

Où la variable  $g2$  est une variable polytomique qui prend les valeurs suivantes :

- ' ' si l'individu n'a pas d'emploi. (Valeur manquante)
- 1 si le premier emploi est un CDI
- 2 si c'est un CDD
- 3 si c'est un contrat d'intérim
- 4 si c'est un contrat d'apprentissage
- 5 si c'est un contrat de qualification
- 6 si c'est un contrat d'emploi solidarité
- 7 si l'individu est à son propre compte

- 8 si l'individu est employé par un membre de sa famille installé à son propre compte
- 9 sinon.

Ensuite à l'aide de cette variable dichotomique, nous avons réalisé différents tableaux croisés pour voir les relations entre notre variable expliquée cdi et nos variables explicatives. Pour finir, nous utiliserons les outils d'analyse des variables qualitatives pour voir quels sont les effets des variables explicatives et ainsi voir si l'on retrouve ce qui a été observé lors de l'analyse des tableaux croisés.

### 1. Influence du sexe.

Pour commencer, on veut voir si le sexe des individus peut influencer la probabilité d'avoir un CDI comme premier emploi.

Tout d'abord faisons l'étude du tableau croisé entre la variable cdi et la variable sexe qui est une variable dichotomique prenant la valeur 1 si l'individu est un homme et 2 si c'est une femme.

Pourct. en col.	Table de Q1 par cdi			
	Q1(Q1)	Cdi		Total
		0	1	
homme	61.34	64.86		
femme	38.66	35.14		
<b>Total</b>	1901	296	2197	

On observe donc que 64,86% des personnes qui ont eu un CDI comme premier emploi sont des hommes alors que 35,14 % de ces mêmes personnes sont des femmes. De même, 61,34% des personnes ayant un type de contrat autre qu'un CDI comme premier emploi ou n'ayant pas répondu à la question (g2) sont des hommes et 38,66% sont des femmes. Cette constatation est le résultat d'une différence de proportion dans l'échantillon considéré. En effet, la proportion d'homme dans l'échantillon est supérieure à celle de femmes ce qui peut expliquer cette différence de répartition des CDI suivant le sexe.

Maintenant nous allons utiliser l'approche économétrique pour voir l'influence de la variable sexe sur la probabilité d'avoir un CDI comme premier emploi.

modèle économétrique	Logistique	
significativité du modèle à 5%	non	p-value
Test sur les paramètres:LR	1.3637	0.2429
variables explicatives	coefficients	p-value
<b>Sexe</b>		
Homme	référence	référence
Femme	-0.0758	0.2454
Constante	-1.8797	<.0001

Le modèle n'est pas globalement significatif au sens que pour les trois tests de significativité globale de la régression (test de WALD, test de LM, et test LR ci-dessus) on

accepte l'hypothèse de nullité simultanée des coefficients de la régression. Le fait d'être une femme n'a aucune influence sur la probabilité d'avoir un CDI comme premier emploi comparativement au fait d'être un homme.

## 2. Influence du statut marital

Nous allons maintenant essayer de voir si le fait d'être célibataire ou non influence cette probabilité. Pour cela, nous allons commencer par étudier le tableau croisé entre la variable dichotomique q7 (qui vaut 1 si l'individu est célibataire, divorcé ou veuf et 2 s'il est marié ou s'il vit en concubinage) et notre variable cdi.

Pourct. En col.	Table de Q7 par cdi			
	Q7(Q7)	cdi		Total
		0	1	
	<b>célibataire, divorcé(e) ou veuf (ve)</b>	73.96	68.24	
	<b>marié(e), concubin(e)</b>	26.04	31.76	
	<b>Total</b>	1901	296	2197

On peut observer que 68,24% des CDI en premier emploi sont occupés par des célibataires et seulement 31% environ par des personnes mariées. De même, on constate la même tendance pour les autres types de contrats.

Cependant les statistiques montrent que lorsqu'on se place dans la colonne cdi=0 pour les célibataires le pourcentage est plus élevé que dans la colonne cdi=1; alors que pour les mariés dans la colonne cdi=0 le pourcentage est plus faible que dans la colonne cdi=1.

On peut penser que les mariés ont une probabilité plus importante vis-à-vis des CDI que les célibataires. Le tableau ci-dessous explicite cela :

Pourct. en col.	Table de cdi par Q7			
	cdi	Q7(Q7)		Total
		<b>célibataire, ou veuf (ve)</b>	<b>divorcé(e)</b>	
	<b>0</b>	87.44		84.04
	<b>1</b>	12.56		15.96
	<b>Total</b>	1608	589	2197

En effet on remarque que le pourcentage des mariés qui ont un CDI est plus élevé que celui des célibataires. On suppose que le fait d'être marié aura une plus grande influence que le fait d'être célibataire sur la probabilité d'avoir un CDI pour notre premier emploi.

Nous allons vérifier cela économétriquement :

modèle économétrique	Logistique	
significativité du modèle à 5%	oui	p-value
Test significativité globale:LR	4.1416	0.0418
variables explicatives	coefficients	p-value
<b>situation maritale</b>		
marié(e) ou concubin€	référence	référence
célibataire, divorcé€ ou veuf(ve)	-0.1395	0.0392
Constante	-1.8007	<.0001

Tout d'abord voyons si la régression est globalement significative. On utilise les trois tests (ici on se contentera du test LR) on remarque que l'on rejette l'hypothèse nulle à 5% qui est la nullité simultanée de tous les coefficients de la régression on a donc une régression significative. Comme on se trouve dans le cas d'une régression à une variable explicative, on a donc significativité du coefficient devant la variable célibataire. En regardant le signe du coefficient, on remarque que celui-ci est négatif donc le fait d'être célibataire diminue la probabilité d'avoir un CDI lors du premier emploi comparativement au fait d'être marié.

### 3. Influence de la nationalité.

Regardons l'effet de la nationalité sur la probabilité d'obtenir un CDI pour le premier emploi. La variable représentant la nationalité (q5) est une variable dichotomique prenant comme valeur 1 si la personne est de nationalité française et 2 si la personne est de nationalité étrangère.

Le tableau croisé suivant montre que l'on a 97,64% des CDI (premier emploi) détenus par des personnes ayant la nationalité française et seulement 2,36% par des étrangers. Ce qui est logique au regard de la composition de notre base de données (les personnes de nationalité française étant largement surreprésentées dans notre base). Ce que l'on peut remarquer dans un second temps, est que dans la colonne cdi=0 le pourcentage pour les français est plus élevé que dans la colonne cdi=1 (98,58% → 97,64%) ce qui n'est pas le cas pour les personnes de nationalité étrangère. (La proportion varie en faveur des étrangers)

Pourct. en col.	Table de Q5 par cdi			
	Q5(Q5)	cdi		Total
		0	1	
<b>Française</b>	98.58	97.64		
<b>Etrangère</b>	1.42	2.36		
<b>Total</b>	1901	296	2197	

Le tableau ci-dessous nous montre que seulement 13,36% des français obtiennent un CDI lors de leur premier emploi alors que 20,59% des étrangers en obtienne un.

On peut formuler l'hypothèse que le fait d'être étranger doit jouer sur la probabilité d'avoir un CDI.

Pourct. en col.	Table de cdi par Q5			
	Cdi	Q5(Q5)		Total
		française	étrangère	
0	86.64	79.41		
1	13.36	20.59		
<b>Total</b>	2163	34	2197	

Malgré nos attentes, nous constatons que notre régression n'est pas significative à 5%. Donc il n'est pas pertinent de prendre comme variable explicative la nationalité comme seule variable exogène.

Modèle économétrique	Logistique	
Significativité du modèle à 5%	non	p-value
Test significativité globale:LR	1.3260	0.2495
variables explicatives	coefficients	p-value
<b>Nationalité</b>		
Etrangère	référence	référence
Française	-0.2597	0.2257
Constante	-1.6097	<.0001

#### 4. Influence de la taille de l'entreprise.

Etudions maintenant l'influence de la taille de l'entreprise sur le type de notre premier contrat. La taille de l'entreprise est modélisée par une variable polytomique (g3) qui prend les valeurs suivantes : 1 si l'entreprise n'a aucun salarié, 2 si elle a moins de 10 salariés, 3 si elle a entre 10 et 49 salariés, 4 si elle a entre 50 et 99 salariés, 5 si elle a entre 100 et 199 salariés, 6 si elle a 200 salariés et plus et 7 si le nombre de salariés de l'entreprise est inconnu.

Pourct. en col.	Table de G3 par cdi			
	G3(G3)	cdi		Total
		0	1	
aucun	1.23	0.68		
moins de 10	22.80	37.84		

<b>10 à 49</b>	26.19	33.78	
<b>50 à 99</b>	10.63	7.09	
<b>100 à 199</b>	7.70	5.41	
<b>200 et plus</b>	20.49	12.50	
<b>inconnu</b>	10.94	2.70	
<b>Total</b>	649	296	945
<b>Fréquence manquante = 1252</b>			

Dans le tableau précédent, les PME de 0 à 50 employés regroupent environ 71,62% des CDI obtenus en premier emploi. De même 25% des CDI en premier emploi se sont trouvés dans les grandes entreprises ayant plus de 50 employés. On peut penser que la taille de l'entreprise peut avoir un impact sur la probabilité d'obtenir un CDI comme premier contrat. Ainsi on peut dire que plus la taille des entreprises diminue et plus la proportion des CDI augmente et peut être plus on a de chance de se voir proposer un CDI.

Dans le tableau précédent, le nombre de fréquence manquante correspond aux nombres d'individus n'ayant pas répondu à la question.

Le tableau ci-dessous illustre bien cette idée ; en effet dans la colonne « moins de 10 employés » la proportion des CDI pour un premier emploi dans cette catégorie d'entreprise par rapport aux autres types de contrat est quasi équivalente ( $43,08\% \approx 56,92\%$ ). Cette caractéristique ne se retrouve pas chez les entreprises de plus de 50 employés. Au sens où plus le nombre de salariés dans l'entreprise augmente et plus la différence entre les deux proportions définies précédemment augmente.

<b>Pourct. en col.</b>	<b>Table de cdi par G3</b>							
	<b>cdi</b>	<b>G3(G3)</b>						<b>Total</b>
		<b>aucun</b>	<b>moins de 10</b>	<b>10 à 49</b>	<b>50 à 99</b>	<b>100 à 199</b>	<b>200 et plus</b>	
<b>0</b>	80.00	56.92	62.96	76.67	75.76	78.24	89.87	
<b>1</b>	20.00	43.08	37.04	23.33	24.24	21.76	10.13	
<b>Total</b>	10	260	270	90	66	170	79	945
<b>Fréquence manquante = 1252</b>								

Etudions économétriquement cela.

modèle économétrique	Logistique	
significativité du modèle à 5%	oui	p-value
Test significativité globale:LR	52.7186	<.0001
<b>taille de l'entreprise</b>		
200 salariés et plus	référence	référence
Aucun	-0.2452	0.7195
10 à 49 salariés	0.6104	0.0006
50 à 99 salariés	-0.0485	0.8482
100 à 199 salariés	0.00161	0.9954
moins de 10 salariés	0.8623	<.0001
Inconnu	-1.0422	0.0025
Constante	-1.1410	<.0001

Tout d'abord on peut constater que notre modèle est globalement significatif car on rejette l'hypothèse nulle à 5%. En ce qui concerne les variables, on remarque qu'à 5% on accepte  $H_0$  pour toutes les variables sauf pour la constante et les variables « 10 à 49 salariés », « moins de 10 salariés » et « inconnu ». Toutes ces variables sauf les 3 significatives peuvent être retirées du modèle. En ce qui concerne les 3 variables significatives, on peut maintenant étudier leurs signes. Pour les variables « 10 à 49 salariés » et « moins de 10 salariés » le signe du coefficient est positif donc le fait d'être dans une entreprise de 10 à 49 salariés ou de moins de 10 salariés augmente la probabilité d'avoir un CDI en premier emploi comparativement au fait d'appartenir à une entreprise de « 200 salariés et plus ». Par contre le signe du coefficient devant « inconnu » est négatif donc le fait d'être dans une entreprise dont le nombre de salariés est « inconnu » diminue la probabilité d'avoir un CDI en sortant du système scolaire comparativement au fait d'appartenir à une entreprise de plus de 200 salariés.

## 5. Influence du type d'intermédiaire

Nous allons maintenant voir s'il est préférable de passer par l'ANPE, par une agence d'intérim,... afin de trouver un CDI lorsque l'on sort du système scolaire. En gros, on veut voir l'efficacité du réseau d'information sur le type de premier emploi c'est-à-dire si le fait de passer par l'ANPE pour faire sa recherche d'emploi par exemple augmente la probabilité d'avoir un CDI. Pour ce faire, on étudie les tableaux croisés entre les variables cdi et g12. La variable g12 est la variable représentant l'intermédiaire par lequel l'individu est passé pour obtenir son premier contrat. C'est une variable dichotomique qui prend les valeurs suivantes : 1 s'il est passé par l'ANPE, 2 s'il est passé par l'intermédiaire d'une agence d'intérim, 3 s'il a répondu à une annonce parue dans la presse ou sur Internet, 4 s'il est passé par l'intermédiaire d'une association d'ancien élève, 5 s'il est passé par son établissement scolaire, 6 s'il est passé par une de ces relations, 7 sinon.

Pourct. en col.

Table de G12 par cdi			
G12(G12)	cdi		Total
	0	1	
ANPE	8.02	9.22	
agence intérim	36.73	4.44	
annonce presse ou sur internet	9.26	16.72	
A son établissement scolaire	7.25	18.77	
par relation	30.40	35.49	
autres	8.33	15.36	
<b>Total</b>	<b>648</b>	<b>293</b>	<b>941</b>
<b>Fréquence manquante = 1256</b>			

En étudiant le tableau précédent, on se rend compte la meilleure façon d'avoir un CDI est de passer par ses relations. En effet, les CDI trouvés grâce aux relations représentent environ 35% de l'ensemble des CDI. Ensuite, le deuxième meilleur moyen est de passer par son établissement scolaire (19% des CDI). Puis viennent les annonces par Internet ou dans la presse (17%) et les autres démarches (16%). Et enfin viennent l'ANPE (9%) et les agences d'intérim (4%).

Maintenant voyons l'étude économétrique de la relation unissant les intermédiaires et la variable cdi :

modèle économétrique	Logistique	
significativité du modèle à 5%	oui	p-value
Test significativité globale:LR	148.1908	<.0001
variables explicatives	coefficients	p-value
intermédiaire		
établissement scolaire	référence	Référence
ANPE	0.0827	0.6965
Intérim	-2.1688	<.0001
annonce presse ou Internet	0.5356	0.0028
par relation	0.0993	0.4498
Autres	0.5558	0.0028
Constante	-0.7381	<.0001

Tout d'abord nous pouvons voir que la régression a un sens car elle est globalement significative à 5%. En ce qui concerne les coefficients devant les variables, on peut voir que seuls les coefficients devant intérim, annonce presse ou Internet et autres sont les seuls que l'on peut considérer significatifs. Discutons du signe des coefficients de ces variables : le signe des coefficients devant annonce presse ou Internet et autres sont positifs cela veut dire que le fait de passer par ces intermédiaires va augmenter la probabilité d'avoir un CDI à la sortie du système scolaire comparativement au fait de passer par son établissement scolaire. Par contre le signe du coefficient devant intérim est négatif ce qui veut dire que passer par une agence d'intérim diminue la probabilité d'avoir un CDI comme premier emploi relativement au fait de passer par son établissement scolaire.

## **6. Influence du secteur d'activité :**

Le secteur d'activité est modélisé par une variable (g9) polytomique qui prend les valeurs suivantes :

- 1 si agriculture**
- 2 si bâtiment du gros oeuvre**
- 3 si bâtiment du second degré**
- 4 si industrie mécanique, métallurgie**
- 5 si industrie agroalimentaire**
- 6 si autres industries (bois, textile, chimie, graphique,..)**
- 7 si transport**
- 8 si immobilier, commerce**
- 9 si réparation automobile**
- 10 si hôtellerie, restauration**
- 11 si administration ou collectivité territoriale**
- 12 si enseignement, santé**
- 13 si services aux particuliers**
- 14 si services aux entreprises**
- 15 si autres**

Ci dessous le tableau croisé de G9 par cdi avec des pourcentages en colonne:

<b>Table de G9 par cdi</b>			
<b>G9(G9)</b>	<b>cdi</b>		<b>Total</b>
	<b>0</b>	<b>1</b>	
	<b>Freq / %</b>	<b>Freq / %</b>	<b>Fréquence</b>
<b>agriculture</b>	3.40 4.94	1.28 4.10	4.68
<b>bâtiment du gros oeuvre</b>	3.51 5.09	2.23 7.17	5.74
<b>Bâtiment du second degré</b>	3.72 5.40	2.98 9.56	6.70
<b>Industrie mécanique, métallurgie</b>	9.67 14.04	1.81 5.80	11.48
<b>Industrie agroalimentaire</b>	5.74 8.33	3.40 10.92	9.14
<b>autres industries (bois, textile, chimie, graphique,..)</b>	9.35 13.58	2.55 8.19	11.90
<b>transport</b>	1.81 2.62	0.64 2.05	2.44
<b>immobilier, commerce</b>	7.12 10.34	2.98 9.56	10.10
<b>réparation automobile</b>	1.81 2.62	1.59 5.12	3.40
<b>hôtellerie, restauration</b>	4.25 6.17	5.63 18.09	9.88
<b>administration ou collectivité territoriale</b>	2.87 4.17	0.53 1.71	3.40

enseignement, santé	3.08 4.48	0.64 2.05	3.72
services aux particuliers	4.68 6.79	2.55 8.19	7.23
Services aux entreprises	2.02 2.93	0.53 1.71	2.55
autres	5.84 8.49	1.81 5.80	7.65
Total	648 68.86	293 31.14	941 100.00
Fréquence manquante = 1256			

En **rouge** les secteurs d'activité qui lors du premier emploi distribuent autant (approximativement) de CDI qu'un autre type de contrat.

En **bleu** ce sont les secteurs d'activités où dans notre échantillon les personnes trouvant leur premier emploi sont majoritairement des contrats à longueur indéterminée.

Enfin, en **vert** ce sont les secteurs d'activité distribuant en proportionnellement moins de CDI qu'un autre type de contrat.

On peut peut-être penser que les secteurs d'activités ayant leurs pourcentages sont en **vert** dans le tableau ci-dessus, sont des secteurs où la situation de leurs salariés est plus précaire que dans d'autres secteurs et l'on peut penser qu'il vaut mieux chercher un emploi dans les secteurs en **bleu** pour maximiser ses chances d'avoir un CDI.

Passons à l'étude économétrique:

modèle économétrique	Logistique	
significativité du modèle à 5%	oui	p-value
Test significativité globale:LR	68,6426	<.0001
variables explicatives	coefficients	p-value
<b>Secteur d'activité</b>		
Transport	référence	référence
administration ou collectivité territoriale	-0.8003	0,0829
agriculture	-0.0948	0,7719
autres	-0.2881	0,2907
autres industries (bois, textile, chimie, graphique,..)	-0.4132	0,0741
Bâtiment du gros œuvre	0,4341	0,1132
Bâtiment du second degré	0,6629	0.0084
enseignement, santé	-0.6895	0,106
Hôtellerie, restauration	1,1675	<.0001
immobilier, commerce	0,0136	0,9523
Industrie agroalimentaire	0,3628	0,1072
Industrie mécanique, métallurgie	-0.7916	0.0024
réparation automobile	0,7609	0.0257
services aux particuliers	0,2799	0,2662
Services aux entreprises	-0.4489	0,3456
Constante	-0.8861	<.0001

On a 4 variables significatives et la constante.

Nos 4 variables sont « bâtiment du second degré », « hôtellerie, restauration », « industrie mécanique, métallurgie », et « réparation automobile ».

Ainsi ces 4 secteurs d'activité influencent la probabilité d'avoir un CDI: le fait de travailler dans le secteur « bâtiment du second degré » influence positivement la probabilité d'avoir un CDI comparativement au fait d'être dans les transports (coefficient= 0,66) car le signe du coefficient est positif.

De même l'« hôtellerie, restauration » (coefficient= 1,1675) et la« réparation automobile » (coefficient= 0,7609) influencent positivement la probabilité d'avoir un CDI lors de notre premier emploi comparativement au fait d'être dans les transports.

Enfin le cas de l'« industrie mécanique, métallurgie » est contraire; le coefficient de ce secteur d'activité étant négatif cette variable comparativement à la variable transports influencera négativement la probabilité d'obtention d'un CDI.

Interprétation économétrique mis à part, on peut penser que réellement il est plus difficile pour un jeune diplômé de trouver un CDI dans le domaine de l'industrie mécanique ou métallurgique face à celui des transports. C'est une conséquence probablement des différents niveaux de dynamisme des secteurs d'activité dans la région centre vis à vis de celui des transports.

## **II. Influences des différentes variables explicatives :**

Dans le cadre de l'approche économétrique qui va suivre, notre but est d'ajouter des variables aux variables précédentes afin de former des résultats cohérents, étant le plus significatif possible et ayant un pouvoir de prédiction assez fort. Nous allons donc dans nos modèles suivants introduire des variables qui nous semblent pertinentes et qui nous permettront d'expliquer au mieux quelles sont les influences qui modifient la probabilité d'avoir un CDI comme premier emploi.

Nous allons maintenant prendre en compte l'influence de plusieurs variables sur la probabilité d'avoir un CDI pour premier emploi.

- a) Dans notre premier modèle que l'on nommera modèle 1, on modélisera l'impact des variables sexe, nationalité, situation maritale.
- b) Dans un deuxième modèle, nommé modèle 2, on ajoutera l'influence des variables : taille de l'entreprise, la nature de l'entreprise et le secteur d'activité pour voir si l'entreprise peut jouer un rôle dans la probabilité d'avoir un CDI en premier emploi.
- c) Dans notre troisième modèle que l'on nommera modèle 3, on ajoutera aux variables précédentes : les variables types d'intermédiaire, mode recrutement ce qui nous permettra de voir l'influence du type de démarche que l'individu peut faire pour avoir un emploi.
- d) Dans un quatrième modèle, nommé modèle 4, on ajoutera aux variables du précédent modèle les variables relatives aux professions des parents dont les deux cas la modalité de référence est « cadre », au fait que l'individu ait des enfants ou non. Cette dernière variable étant croisée avec la variable « situation maritale » et la modalité de référence pour cette variable croisée est donc « marié avec au moins un enfant ». Dans ce modèle, on cherchera à voir si la profession des parents et le fait d'avoir au moins un enfant va influencer la probabilité d'avoir un contrat à durée indéterminé comme premier emploi.

- e) Dans un modèle 5, on ajoutera aux variables du modèle 4 la variable « dernier diplôme obtenu » afin de voir si la théorie du signal est vérifiée dans notre cas. Au sens où l'on va essayer de voir si le niveau d'étude a un impact sur la probabilité que l'on souhaite étudier.

### 1) Modèle 1 :

Dans notre premier modèle, notre but est de voir si le sexe de l'individu, sa nationalité ou sa situation maritale peuvent avoir un impact sur la probabilité qu'il ait un CDI en sortant du système scolaire toutes choses étant égales par ailleurs.

Les trois variables explicatives sont toutes des variables dichotomiques définies dans les différents points précédents. L'intérêt d'une telle modélisation est de déterminer les éventuelles discriminations pouvant intervenir dans la probabilité d'avoir un CDI. En effet, la variable sexe nous permettra de mettre en évidence une discrimination en fonction du sexe, la variable nationalité, une discrimination raciale et la variable situation maritale, une discrimination vis-à-vis des célibataires ou des mariés(es).

Pour la variable sexe nous prenons comme modalité de référence le fait que l'individu soit un homme. Pour la variable situation maritale, la modalité de référence est « marié » et pour la variable nationalité, on prend comme modalité de référence la nationalité française.

Au vu des résultats (**cf. : annexe 2**), on peut dire que ce modèle est significatif car on rejette l'hypothèse que tous les coefficients sauf la constante sont nuls à 5% notre modèle a donc un sens et son interprétation vaut le coup d'être faite. A 5%, le coefficient devant la variable femme est non significatif donc le fait d'être une femme n'influence pas la probabilité relativement au fait d'être un homme. De même, le coefficient devant la variable étrangère n'est pas significatif donc le fait d'être de nationalité étrangère n'influence pas la probabilité comparativement au fait d'être de nationalité française. Par contre, le coefficient devant célibataire est significatif (à 5%) et négatif ce qui veut dire que le fait d'être célibataire diminue la probabilité d'avoir un CDI comme premier emploi comparativement au fait d'être marié. On peut faire l'hypothèse d'une certaine forme de discrimination vis-à-vis des célibataires comparativement aux mariés. Par contre, on ne peut faire aucune hypothèse de discrimination en fonction du sexe et de la nationalité.

Il faut noter que ce modèle au même titre que les modèles de la partie précédente sont assez restrictifs dans le sens où l'on peut aisément imaginer que d'autres paramètres interviennent et que, seules, les caractéristiques des individus ne suffisent pas pour déterminer les influences sur la probabilité d'avoir un CDI comme premier emploi. On peut donc raisonnablement imaginer que les caractéristiques des entreprises ont aussi un rôle à jouer.

### 2) Modèle 2 :

Dans ce modèle, on prend en compte les variables du modèle 1 auxquelles l'on ajoute les variables qui décrivent les caractéristiques de l'entreprise dans laquelle le premier emploi a été trouvé.

Parmi ces variables, on a la variable « taille de l'entreprise » qui est une variable polytomique (définie plus haut) qui va nous permettre de voir si les petites entreprises ont plus d'impact que les « grosses » entreprises dans la probabilité. La modalité de référence est « 200 salariés et plus ». On a aussi introduit la variable polytomique « nature de l'entreprise » qui prend les valeurs suivantes :

1 si l'entreprise est une entreprise privée

- 2 si c'est une association
- 3 si c'est une entreprise publique
- 4 si c'est une entreprise de la fonction publique
- 5 sinon

et qui va nous permettre de déterminer l'influence de la nature des entreprises sur la probabilité. La modalité de référence est dans ce cas « entreprise privée ».

De même on a introduit la variable polytomique « secteur d'activité » (définie elle aussi dans une section ci-dessus) pour voir si le secteur d'activité des entreprises jouent un rôle sur notre probabilité. La variable de référence est « transport ».

Regardons maintenant les résultats de notre régression de la variable cdi en fonction des variables définies précédemment (cf. : **annexe 2**).

Tout d'abord on peut remarquer que notre modèle est globalement significatif à 10% donc il existe au moins un coefficients autre que la constante qui soit différent de 0.

Prenons les valeurs de nos paramètres une par une et déterminons ceux qui sont significatifs à 5 %.

De cette étude ressort que les paramètres devant les variables suivantes sont significatifs et ont un signe positif : moins de 10 salariés, entre 10 et 49 salariés, hôtellerie et restauration, service aux particuliers. Les variables précédentes influencent donc positivement la probabilité relativement à leur variable de référence. Ainsi, le fait d'être dans une PME (petite et moyenne entreprise qui a entre 10 et 50 salariés) augmente la probabilité d'avoir un CDI en premier emploi comparativement au fait d'être dans une grande entreprise de plus de 200 salariés. De même, être dans une entreprise qui appartient aux secteurs du service aux particuliers ou à l'hôtellerie et restauration augmente la probabilité d'avoir un CDI en sortant du système scolaire relativement au secteur des transports. On peut donc faire l'hypothèse que l'on a une probabilité d'avoir un CDI plus forte lorsque l'on postule à des emplois dans les PME, plutôt que dans des grandes entreprises, et dans des entreprises appartenant au secteur du service aux particuliers et au secteur de l'hôtellerie et de la restauration plutôt que dans des entreprises du secteur des transports.

Inversement les variables suivantes sont elles aussi significatives mais leur paramètres sont de signe négatif : femme, célibataire, taille de l'entreprise inconnue, industrie métallurgique et mécanique. Ces variables influencent donc négativement la probabilité. Ainsi, le fait d'être une femme diminue la probabilité d'avoir un CDI comme premier emploi comparativement au fait d'être un homme. Ce résultat nous permet de formuler l'hypothèse de présence d'effets discriminatoires vis-à-vis du sexe lors du premier emploi. De même, le fait d'être célibataire diminue la probabilité comparativement au fait d'être marié. De plus, le fait que l'entreprise soit dans le secteur de l'industrie mécanique et métallurgique baisse la probabilité relativement au secteur des transports. On peut donc faire l'hypothèse que l'on a une probabilité plus faible de signer un CDI à la sortie du système scolaire en étant une femme (comparé au fait d'être un homme), en étant célibataire (comparé au fait d'être marié) et en recherchant un emploi dans le secteur de l'industrie mécanique et métallurgique (comparé au secteur des transports).

On peut noter que tous les coefficients de la variable « nature de l'entreprise » sont tous non significatifs ce qui revient à dire que cette variable n'a aucune influence sur la probabilité que l'on désire modéliser.

Il peut être intéressant de compléter notre modèle en prenant en compte les moyens que se servent les individus pour trouver un travail.

### 3) Modèle 3 :

Dans ce modèle, on reprend les variables du modèle 2 auxquelles l'on ajoute deux variables. La première est la variable « intermédiaire pour trouver l'emploi » qui est une variable polytomique définie dans une section plus haut et dont la modalité de référence est « établissement scolaire ». La deuxième variable est la variable polytomique représentant le type de démarche personnelle faite par l'individu pour trouver son emploi qui prend les valeurs suivantes :

- 1 si l'individu a été recruté en écrivant ou en se présentant à l'entreprise
- 2 si l'individu a passé les concours de la fonction publique
- 3 si c'est la suite d'un premier CDD dans l'entreprise
- 4 si c'est la suite d'un stage dans l'entreprise
- 5 si c'est suite à un contrat d'intérim dans l'entreprise
- 6 si c'est suite à un contrat d'apprentissage dans l'entreprise
- 7 sinon

La modalité de référence est « suite à un stage ».

En introduisant ces variables, on cherche à déterminer les meilleures démarches que peut entreprendre un individu à sa sortie du système scolaire afin d'obtenir un CDI.

Étudions les résultats de la régression de CDI en fonction de toutes les variables définies précédemment (**cf. : annexe 2**).

En utilisant la valeur du Likelihood ratio, on peut dire que notre modèle est globalement significatif à 5%. Regardons maintenant quels sont les paramètres significatifs de la régression.

À 5%, les coefficients devant les variables « entre 10 et 49 salariés », « hôtellerie et restauration » et « autres » intermédiaires sont significatifs et positifs. Cela veut dire que le fait d'appartenir à une PME augmente la probabilité d'avoir un CDI comme premier emploi comparativement au fait d'être dans une grande entreprise. De même, le fait d'être dans le secteur hôtellerie et restauration accroît aussi cette probabilité comparativement au fait d'être dans le secteur des transports ou encore le fait d'avoir utilisé un autre type d'intermédiaire que ceux proposés pour prendre connaissance de l'emploi augmente aussi cette probabilité comparativement au fait d'être passé par son établissement scolaire.

Inversement les coefficients devant les variables suivantes sont significatifs et négatifs : femme, célibataire, agence d'intérim. On retrouve les résultats du modèle 2 qui nous disent qu'être une femme et qu'être célibataire diminue la probabilité d'avoir un CDI comme premier emploi. De même, le fait de passer par une agence d'intérim diminue aussi cette probabilité relativement au fait de passer par son établissement scolaire.

On peut supposer que la probabilité d'avoir un CDI est plus forte si l'on postule dans une PME que si l'on brigue un poste dans une grande entreprise. Peut-être que les grandes entreprises ont une politique d'embauche plus méfiante vis-à-vis des jeunes entrant sur le marché du travail. Même chose pour les secteurs d'activité : soit certains secteurs d'activité peuvent présenter moins de méfiance vis-à-vis des jeunes diplômés, soit cette différence d'influence sur la probabilité de trouver un CDI lors de notre premier emploi peut refléter le niveau de dynamisme de cette filière.

En passant par une agence d'intérim, on a une probabilité plus faible d'avoir un CDI qu'en passant par son établissement scolaire (qui est la référence). Cela tient peut-être du fait que les agences d'intérim proposent principalement des contrats dit « précaires » alors que les établissements scolaires, grâce à leurs relations privilégiées avec les entreprises, ont plus de facilités à trouver des CDI.

On peut noter que les coefficients des variables « nature de l'entreprise » et « type de démarche personnelle » sont tous non significatifs ce qui veut dire que toutes les modalités de ces deux variables ne jouent aucun rôle sur la probabilité que l'on désire analyser.

#### 4) Modèle 4 :

On reprend toutes les variables du modèle 3 en rajoutant les variables polytomiques q10 et q11 qui sont la profession du père et de la mère et la variable q8 qui va être croisée avec la variable q7. Les deux variables q10 et q11 prennent différentes modalités : 1 si « agriculteur exploitant », 2 si « artisan, commerçant, chef d'entreprise », 3 si « cadre », 4 si « agent de maîtrise, technicien », 5 si « employé », 6 si « ouvrier », et 7 si « autre ». Voyant que dans la base de données nous avons présence d'une autre modalité inconnue car pas présente dans le questionnaire, et codée 7, nous avons donc rajouté un format assigné à cette dernière sous le nom de « donnée inconnue ». Ainsi ces 2 variables rajoutées dans ce nouveau modèle, prenant pour variable expliquée la variable dichotomique cdi, possède 8 modalités. On prend pour ces deux variables la modalité « cadre » pour référence. La variable q8, quant à elle, a 2 modalités : « avoir au moins un enfant à charge » et « n'avoir pas d'enfant » c'est donc une variable dichotomique. Pour cette variable, on prend la modalité « avoir un enfant à charge » comme référence. Pour des raisons pratiques, nous avons décidé de la croisée avec la variable « situation maritale » et de prendre comme modalité de référence pour cette variable croisée « marié avec au moins un enfant ».

Ce nouveau modèle reste valide : le test LRT montrant la significativité globale de la régression. (cf. : **annexe 3**) Cependant, le nombre de coefficient (à 5%) restant significativement différent de zéro reste relativement faible; nous allons néanmoins expliquer ce que l'on peut déduire du modèle 4 : en regardant uniquement les modalités significatives, on peut dire que comparativement à l'homme, le fait d'être une femme augmente la probabilité d'avoir un CDI lors de son premier emploi (on peut noter que ce résultat est contraire à celui que l'on a pu dégager de l'étude de nos modèles précédents). Idée contraire pour le fait d'être célibataire et de n'avoir pas d'enfant comparativement aux personnes mariés et ayant au moins un enfant. Le fait d'être dans des entreprises de moins de 10 salariés ou de 10 à 49 salariés augmente la probabilité d'avoir un CDI face aux entreprises de plus de 200 salariés. Idée contraire pour la modalité «taille de l'entreprise inconnue » face aux grandes firmes. Pour la variable « secteur d'activité » on peut seulement dire que « l'industrie mécanique, métallurgique » présente une plus faible probabilité d'y obtenir un CDI face au secteur des transports, alors que l'on est dans le cas contraire avec « hôtellerie et restauration » vis-à-vis des transports. Enfin la dernière variable présentant des modalités dont les coefficients sont significatifs est « comment avez-vous pris connaissance de cet emploi » : on peut dire ici que la probabilité d'obtenir un CDI pour son premier emploi est plus faible en passant par l'intermédiaire d'une agence d'intérim que par son établissement scolaire. Idée contraire avec la modalité « autres » vis-à-vis de la modalité « établissement scolaire » puisque son coefficient est significativement différent de zéro d'après le test WALD et sa valeur est positive.

On ne peut cependant rien dire à propos des modalités des variables rajoutées dans ce nouveau modèle puisque tous les coefficients ne sont pas significativement différents de zéro.

En gros ce modèle ne nous apporte rien de plus que le modèle précédent. En effet, on retrouve les mêmes résultats que ceux trouvés précédemment (mis à part le résultat contraire trouver pour le sexe).

Voyons dans un cinquième modèle si le dernier diplôme obtenu par l'individu influence la probabilité d'avoir un CDI comme premier emploi.

### 5) Modèle 5 :

Dans ce modèle, on garde les variables explicatives du modèle 4 et on ajoute en plus la variable « dernier diplôme » qui est une variable polytomique qui prend les valeurs suivantes :

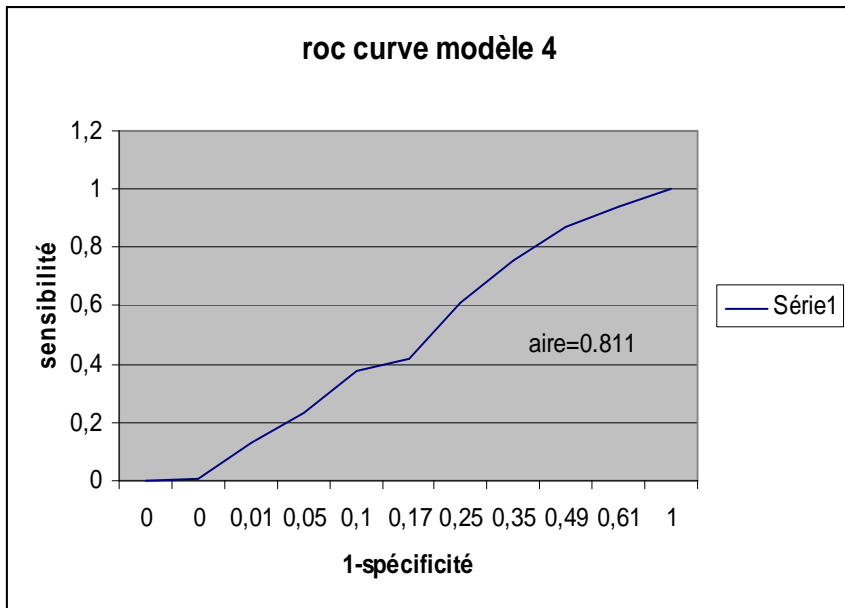
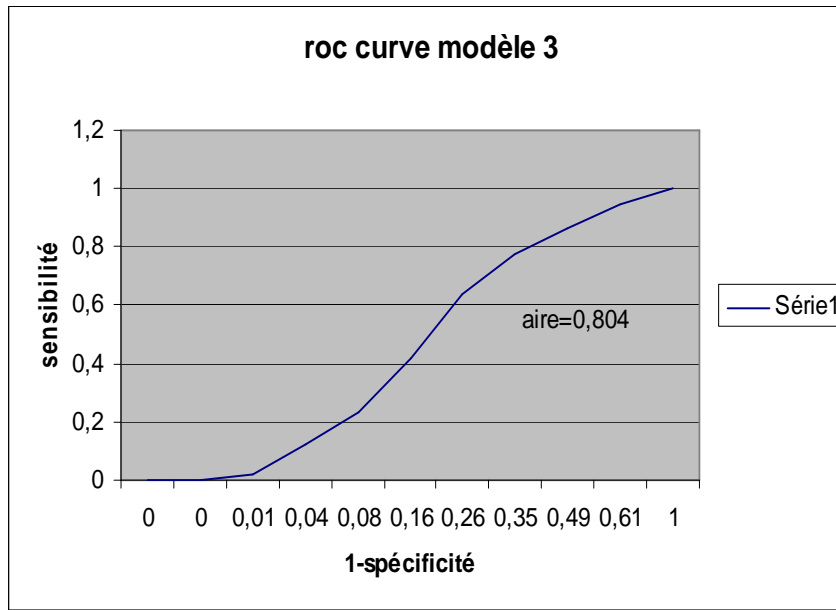
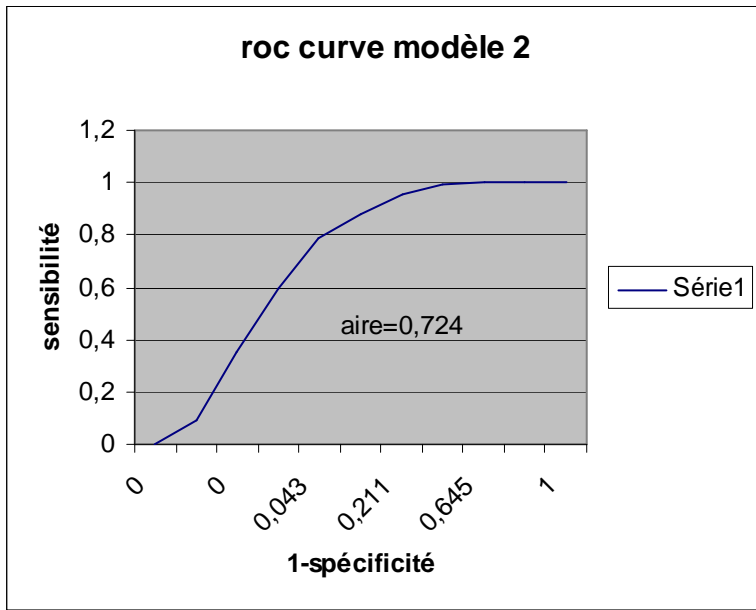
- 01 si brevet
- 02 si CAP
- 03 si BEP
- 04 si mentions complémentaires
- 05 si brevets de technicien
- 06 si brevets de maîtrise
- 07 si brevets professionnels
- 08 si bacs technologiques
- 09 si bacs professionnels
- 10 si bacs généraux
- 11 si BTS
- 12 si DUT
- 13 si DEUG
- 99 autre

Pour cette variable, nous prendrons comme modalité de référence « CAP ».

Le modèle est globalement significatif à 5%. (cf. : **annexe 3**).

Cependant certaines modalités significatives dans les modèles précédents ne le sont plus après l'introduction de cette nouvelle variable. D'ailleurs, cette dernière ne présente aucun coefficient significatif ; donc le dernier diplôme obtenu ne joue aucun rôle sur la probabilité d'obtenir un CDI dans son premier emploi (d'après notre modèle 5).





Les ROC CURVE sont régulièrement utilisés pour mesurer l'efficacité d'un test diagnostic.

On choisit de prendre le modèle qui présente la plus grande aire sous la courbe. On considère qu'un modèle est bon si l'aire est supérieure à 0,5. Ainsi pour un mauvais modèle, l'aire est proche de 0,5 et pour un modèle parfait l'aire est proche de 1. Un autre critère peut jouer sur cette décision ainsi si plus la courbe est concave meilleur est le modèle. Dans notre cas, on peut voir aisément que le premier modèle est un mauvais modèle car l'aire est proche de 0.5 et sa courbe n'est pas très concave. Grâce à nos résultats précédents, on peut écarter le modèle 5 car très peu de variables dans ce modèle jouent un rôle sur la probabilité d'avoir un CDI comme premier emploi et ce modèle n'aurait donc que très peu d'intérêt à être interprété. De plus la seule variable qui est introduite en plus dans ce modèle est non significative donc il n'apporte rien de plus par rapport au modèle 4 même si ce modèle présente la plus grande aire.

Il nous reste donc trois modèles parmi lesquels il faut en choisir un. On va choisir le modèle avec la plus grande aire sous la courbe. Le modèle 2 est celui possédant l'aire la plus faible donc on peut l'écarter. Par contre les modèles 3 et 4 ont une aire assez proche même si le 4 a une aire légèrement supérieure au 3. Ce critère ne nous permet pas de départager les deux modèles.

Un autre critère peut nous permettre de choisir : en utilisant le critère AIC.

<b>Statistiques d'ajustement du modèle 3</b>		
<b>Critère</b>	<b>Coordonnée à l'origine uniquement</b>	<b>Coordonnée à l'origine et covariables</b>
<b>AIC</b>	1169.203	990.851
<b>SC</b>	1174.050	1179.882

Statistiques d'ajustement du modèle 4		
Critère	Coordonnée à l'origine uniquement	Coordonnée à l'origine et covariables
AIC	1169.203	1009.099
SC	1174.050	1265.987

On choisit le modèle qui présente un critère AIC le plus faible. Dans notre cas, on retient le modèle 3. Ce choix peut se confirmer si l'on regarde les résultats des modèles. En effet, l'ajout des variables relatifs aux professions des parents dans le modèle 4 n'apporte rien de plus par rapport au modèle 3 car les coefficients devant ces variables ne sont pas significatives. Le modèle que l'on retient donc est le modèle 3 qui a les caractéristiques suivantes :

Variable expliquée : cdi.

Variables explicatives : le sexe, la nationalité, la situation maritale, la taille de l'entreprise, la nature de l'entreprise, le secteur d'activité, les intermédiaires qui ont permis la prise de connaissance de l'emploi, la démarche personnelle.

Nous allons maintenant calculer les effets marginaux des modalités des variables qui sont significatives pour le modèle 3. A savoir : femme, célibataire, entre 10 et 49 salariés, hôtellerie et restauration, agence d'intérim, autres intermédiaires.

On prend comme profil type : une femme célibataire qui a trouvé son premier emploi dans une entreprise d'« hôtellerie restauration » qui comprend « entre 10 et 49 salariés » « par l'intermédiaire d'une agence intérim ». On obtient ce tableau que l'on nommera : **Tableau 1.**

profil déterminant l'index	Index	proba:[cdi=1]	cote	Variables	effets marginaux	élasticité
femme célibataire travaillant dans "l'hôtellerie restauration"	-1,967	0,12273	0,14	Femme	-0,03256175	-0,26528228
une entreprise "de 10 à 49 salariés"	-1,967	0,12273	0,14	Célibataire	-0,02543348	-0,20720792
trouvé par l'intermédiaire d'une agence d'intérim	-1,967	0,12273	0,14	"hôtellerie restauration"	0,08201814	0,66820606
	-1,967	0,12273	0,14	"de 10 à 49 salariés"	0,04759356	0,38774725
	-1,967	0,12273	0,14	Agence intérim	-0,21099454	-1,71898356

Ce tableau va nous servir de référence afin de calculer les déformations de probabilités par rapport au profil moyen induite par une modification des modalités de ce profil.

La cote représente le rapport de la probabilité associée à l'événement  $c_{di}=1$  à la probabilité de non survenue de cet événement; la cote est ici égale à 0,14. Ainsi on conclut que l'individu ayant le profil du tableau ci-dessus a 0,14 fois plus de chance d'obtenir un CDI dès son premier emploi que de ne pas en obtenir un.

Pour déterminer l'influence du sexe sur la probabilité, on va maintenant remplacer femme par homme dans le profil type. On obtient les résultats suivants :

profil déterminant l'index	index	proba:[c <sub>di</sub> =1]	cote	variables	effets marginaux	élasticité
homme célibataire travaillant dans "l'hôtellerie restauration" une entreprise "de 10 à 49 salariés" trouvé par l'intermédiaire d'une agence d'intérim	-1,3619	0,2039	0,26	Homme	0,04909269	0,240731061
	-1,3619	0,2039	0,26	célibataire	-0,03834555	-0,188031338
	-1,3619	0,2039	0,26	"hôtellerie restauration"	0,12365708	0,606365242
	-1,3619	0,2039	0,26	"de 10 à 49 salariés"	0,07175585	0,351862199
	-1,3619	0,2039	0,26	agence intérim	-0,31811218	-1,559895881

Première comparaison que l'on peut faire par rapport au tableau 1, tout d'abord on peut remarquer que la probabilité d'avoir un CDI comme premier emploi est plus forte dans le profil avec l'homme que dans celui avec la femme. De plus, la côte dans le deuxième tableau est là aussi plus importante ainsi l'individu du deuxième profil a 0,26 fois plus de chance d'avoir un CDI comme premier emploi que d'avoir un autre type de contrat ou pas de contrat du tout. En conclusion de cette comparaison, on peut dire que, toutes choses égales par ailleurs, un homme à une plus forte probabilité d'avoir un CDI comme premier emploi qu'une femme toutes choses étant égales par ailleurs. On peut mettre un bémol à ce résultat vu que la différence de probabilité n'est que de 0,08 ce qui n'est pas énorme.

Maintenant, nous reprenons le profil type du début en modifiant la situation maritale. Notre individu célibataire devient marié. On obtient le tableau suivant :

profil déterminant l'index	index	proba:[c <sub>di</sub> =1]	cote	variables	effets marginaux	élasticité
femme mariée travaillant dans "l'hôtellerie restauration" une entreprise "de 10 à 49 salariés" trouvé par l'intermédiaire d'une agence d'intérim	-1,4943	0,18	0,22	Femme	-0,04526525	-0,246976975
	-1,4943	0,18	0,22	Mariée	0,03535599	0,192909926
	-1,4943	0,18	0,22	"hôtellerie restauration"	0,11401633	0,622097758
	-1,4943	0,18	0,22	"de 10 à 49 salariés"	0,06616151	0,360991479
	-1,4943	0,18	0,22	agence intérim	-0,29331102	-1,600368331

Comparons maintenant ces résultats par rapport au tableau 1. On observe une augmentation de la probabilité d'avoir un CDI comme premier emploi et de la cote par rapport aux résultats obtenus en prenant le profil type. On peut donc dire que, toutes choses égales par ailleurs, le fait de changer la situation maritale de notre individu en « marié » a fait croître la probabilité. On peut donc faire l'hypothèse qu'un individu marié aura une plus grande probabilité qu'un individu célibataire toutes choses étant égales par ailleurs. Là aussi, cette différence de probabilité est assez minime.

Nous allons maintenant voir si une modification de la taille dans laquelle les individus postulent à un impact sur la probabilité. Pour cela, nous allons modifier la taille de l'entreprise dans notre profil type. On prend maintenant le cas des entreprises ayant entre 100 et 199 salariés. Comparons les résultats suivants au tableau 1 :

profil déterminant l'index	index	proba:[cdi=1]	cote	Variables	effets marginaux	élasticité
femme célibataire travaillant dans "l'hôtellerie restauration" une entreprise "de 100 à 199 salariés"	-1,9442	0,125	0,14	Femme	-0,03311744	-0,264543401
	-1,9442	0,125	0,14	Célibataire	-0,02586752	-0,206630791
	-1,9442	0,125	0,14	"hôtellerie restauration"	0,08341783	0,666344934
	-1,9442	0,125	0,14	"de 100 à 199 salariés"	0,05091368	0,406700486
trouvé par l'intermédiaire d'une agence d'intérim	-1,9442	0,125	0,14	agence intérim	-0,21459531	-1,714195745

On remarque que les probabilités et les côtes sont assez proches ce qui traduit que contrairement à nos attentes, le fait de modifier la taille de l'entreprise ne joue à aucun impact sur la probabilité d'avoir un CDI en sortant du système scolaire. On ne peut donc pas dire que plus la taille de l'entreprise augmente et plus on a de chance d'avoir un CDI toutes choses égales par ailleurs.

Considérons maintenant une modification du secteur d'activité dans le profil type : prenons par exemple le secteur de l'industrie mécanique et métallurgique.

profil déterminant l'index	index	proba:[cdi=1]	cote	variables	effets marginaux	élasticité
femme célibataire travaillant dans "industrie mécanique..." une entreprise "de 10 à 49 salariés"	-3,32	0,035	0,04	Femme	-0,01018302	-0,291848838
	-3,32	0,035	0,04	célibataire	-0,0079538	-0,227958649
	-3,32	0,035	0,04	"industrie mécanique..."	-0,02007981	-0,575494253
	-3,32	0,035	0,04	"de 10 à 49 salariés"	0,01488391	0,426577997
trouvé par l'intermédiaire d'une agence d'intérim	-3,32	0,035	0,04	agence intérim	-0,0659842	-1,891130285

On observe une très forte baisse de la probabilité et de la cote par rapport aux résultats du profil type. Cela veut dire dans le secteur de l'industrie mécanique et métallurgique la probabilité d'avoir un CDI est plus faible que dans le secteur de l'hôtellerie et de la restauration (toutes choses étant égales par ailleurs). Grâce à ce résultat, on peut formuler l'hypothèse que la probabilité (cdi=1) est différente pour tous les secteurs.

Enfin, si le fait de changer d'intermédiaire pour trouver un emploi fait varier notre probabilité. Ainsi maintenant on se place dans le cas où notre individu du profil type utilise un autre type d'intermédiaire que ceux proposés par le questionnaire.

Profil déterminant l'index	index	proba:[cdi=1]	cote	variables	effets marginaux	élasticité
femme célibataire travaillant	0,7148	0,6715	2	Femme	-0,06670979	-0,099350219
dans "l'hôtellerie restauration"	0,7148	0,6715	2	célibataire	-0,05210599	-0,077600932
une entreprise "de 10 à 49 salariés"	0,7148	0,6715	2	"hôtellerie restauration"	0,1680319	0,250248221
trouvé par l'intermédiaire	0,7148	0,6715	2	"de 10 à 49 salariés"	0,09750571	0,145214275
de "autres"	0,7148	0,6715	2	"autres"	0,16148005	0,24049061

On voit clairement que la probabilité d'avoir un CDI explose par rapport à celle du profil type (même constat pour la cote) elle est 5 fois plus élevée et la cote est 14 fois plus forte. Ainsi l'individu de ce profil a 14 fois plus de chance en plus que l'individu du profil type d'avoir un CDI à sa sortie du système scolaire et il a deux fois plus de chance d'avoir un CDI qu'un autre type de contrat ou pas de contrat du tout.

Voyons maintenant les résultats que l'on obtient si notre individu passe par l'ANPE.

profil déterminant l'index	index	proba:[cdi=1]	cote	Variables	effets marginaux	élasticité
femme célibataire travaillant	-0,2004	0,45	0,8	Femme	-0,07484602	-0,16629974
dans "l'hôtellerie restauration"	-0,2004	0,45	0,8	Célibataire	-0,05846108	-0,129894175
une entreprise "de 10 à 49 salariés"	-0,2004	0,45	0,8	"hôtellerie restauration"	0,18852585	0,418883968
trouvé par l'intermédiaire	-0,2004	0,45	0,8	"de 10 à 49 salariés"	0,10939796	0,243070387
de l'ANPE	-0,2004	0,45	0,8	ANPE	0,04534323	0,100747726

On observe que là aussi la probabilité et la cote ont augmenté. Ainsi, dans ce cas, en passant par l'ANPE, notre individu aura une fois plus de chance d'avoir un CDI toutes choses égales par ailleurs. Au vu des deux derniers résultats, on peut faire l'hypothèse que cette déformation de probabilité imputable au changement d'intermédiaire vient du fait que l'on a plus de chance d'avoir un CDI en passant par l'ANPE (ou par un autre type d'intermédiaire que ceux proposés) qu'en passant par une agence d'intérim (toutes choses étant égales par ailleurs). Hypothèse tout a fait raisonnable car comme son nom l'indique une agence intérim propose principalement des contrats à durée déterminée et non des contrats à durée indéterminée.

## Conclusion :

La principale préoccupation des jeunes dans le système scolaire est de savoir s'ils vont trouver un emploi à la fin de leurs études et quel type de contrat ils vont décrocher. Il est bien évident que le but ultime de tout travailleur est de décrocher un contrat à durée indéterminée. C'est dans cette optique que l'on a choisit à partir de notre base de données de voir comment évoluait la probabilité d'obtenir un CDI comme premier emploi chez les jeunes sortant du système scolaire en fonction de plusieurs caractéristiques qui sont propres à chaque individu et propre à chaque entreprise sur le marché.

Tout au long de ce papier, on a pu se rendre compte que certaines caractéristiques influençaient négativement la probabilité. Ainsi le fait d'être une femme dans l'échantillon considéré fait baisser cette probabilité comparativement au fait d'être un homme.

De même, un individu étant célibataire a moins de chance d'avoir un CDI comme premier emploi qu'une personne qui est mariée.

On a vu aussi que passer par une agence d'intérim pour trouver son premier emploi faisait baisser cette probabilité relativement au fait de passer par son établissement scolaire et qu'il est plus préférable de passer par l'ANPE par exemple que par les agences d'intérim.

Par contre, on a dégagé l'idée selon laquelle le secteur de l'hôtellerie et restauration est un secteur dans lequel on a une probabilité plus forte d'avoir un CDI comme premier emploi comparativement au secteur des transports. Puis nous avons dégagé l'idée selon laquelle d'un secteur à l'autre la probabilité est différente et qu'il peut exister des secteurs « privilégiés ».

Enfin, nous avons vu qu'il est plus judicieux de postuler dans des entreprises de 10 à 49 salariés plutôt que dans les grandes de plus de 200 salariés car elles présentent une probabilité plus forte d'obtention d'un CDI. Idée cependant non vérifiée avec le calcul des déformations de probabilités puisque l'on a vu que de passer d'une taille à l'autre laisse inchangée la probabilité.

Il faut cependant mettre un bémol à toute cette analyse car les augmentations et les diminutions de probabilités sont assez faibles donc les variables explicatives n'ont pas d'impacts très importants sur la probabilité d'avoir un CDI comme premier emploi et peut être que le peu d'impact observé est le résultat des différences de fréquences des modalités pour chaque variable.

Enfin de notre étude, il ressort que pour avoir la plus forte probabilité d'avoir un CDI il faut être : un homme marié, poser sa candidature dans une entreprise ayant entre 10 et 49 salariés par l'intermédiaire de son établissement scolaire et dans le secteur de l'hôtellerie restauration et ne surtout pas passer par une agence d'intérim sous peine de voir cette probabilité fortement diminuée. Mais bien sur, les différences de probabilité entre les profils des individus ne peuvent être entièrement expliquées par ces quelques faits et il y a certainement d'autres paramètres qui doivent être pris en compte tel que le physique, le comportement des individus lors des entretiens, leur motivation pour trouver un emploi, l'adéquation du diplôme avec les besoins des entreprises ...

## **ANNEXE 1 :**

Statistiques descriptives des variables	pourcentage
<b>caractéristiques individuelles</b>	
sexe féminin	38,19%
sexe masculin	61,81%
nationalité française	98,45%
nationalité étrangère	1,55%
Célibataire	73,19%
Marié	26,81%
emploi actuel	75,33%
chômage actuel	24,67%
<b>caractéristiques liées au premier emploi occupé</b>	
<b>contrat:</b>	
CDI	31,32%
CDD	37,78%
contrat d'intérim	26,03%
contrat d'apprentissage	1,38%
contrat emploi jeune	1,80%
<b>taille de l'entreprise:</b>	
Aucun	1,06%
moins de 10 salariés	27,51%
10 à 49 salariés	28,57%
50 à 99 salariés	9,52%
100 à 199 salariés	6,98%
200 salariés et plus	17,99%
?????	8,36%
<b>Secteur d'activité:</b>	
Agriculture	4,68%
bâtiment du gros œuvre, travaux publics	5,74%
bâtiment du second degré (menuiserie, ...)	6,70%
industrie mécanique, métallurgie	11,48%
industrie agro-alimentaire	9,14%
autres industries	11,90%
transport (routier, ferroviaire, maritime, ...)	2,44%
immobilier, commerce	10,10%
réparation automobile	3,40%
hôtellerie restauration	9,88%
administration ou collectivité territoriale	3,40%
enseignement, santé	3,72%
services aux particuliers	7,23%
services aux entreprises	2,55%
autres	7,65%
<b>nature de l'entreprise:</b>	
entreprise privée	84,59%
Association	2,02%
entreprise publique (EDF, SNCF, Air France,...)	7,23%
fonction publique (Etat, collectivités locales, ...)	5,63%
Autres	0,53%
<b>mode de recrutement:</b>	
démarche personnelle	66,10%
<b>nature du temps de travail:</b>	
temps plein	83,26%
temps partiel	16,74%

**ANNEXE 2 :**

modèles économétriques Logit		modèle 1		modèle 2		modèle 3	
références	variables actives	coefficients	p-value	coefficients	p-value	coefficients	p-value
<b>homme</b>	<b>femme</b>	-0.1094	0.1048	-0.3232	0.0005	-0.3024	0.0023
<b>marié</b>	<b>célibataire</b>	-0.1677	0.0163	-0.2635	0.0027	-0.2362	0.0129
<b>français</b>	<b>étrangère</b>	0.2473	0.2502	0.4802	0.0824	0.4226	0.1565
	<b>constante</b>	-1.5794	<.0001	-0.8307	0.0388	-0.6723	0.1413
<b>200 salariés et plus</b>	<b>aucun</b>			-0.4015	0.5740	-1.0206	0.1737
	<b>moins de 10 salariés</b>			0.7029	0.0002	0.3601	0.0754
	<b>entre 10 et 49 salariés</b>			0.6266	0.0010	0.4420	0.0298
	<b>entre 50 et 99 salariés</b>			0.0601	0.8224	0.2701	0.3517
	<b>entre 100 salariés et 199</b>			0.1650	0.5762	0.4649	0.1540
	<b>inconnu</b>			-1.1642	0.0019	-0.6935	0.0735
<b>t r a n s p o r t</b>	<b>administration</b>			-0.3395	0.5269	-0.5052	0.3746
	<b>agriculture</b>			-0.3349	0.3303	-0.4059	0.2544
	<b>autres industries</b>			-0.3663	0.1341	0.00946	0.9722
	<b>bâtiment du gros œuvre</b>			0.0946	0.7471	0.1745	0.5820
	<b>bâtiment du second degré</b>			0.4455	0.1017	0.4628	0.1215
	<b>enseignement, santé</b>			-0.3277	0.4740	-0.6073	0.1989
	<b>hôtellerie restauration</b>			1.0595	<.0001	0.7617	0.0014
	<b>immobilier commerce</b>			0.0623	0.7987	-0.1854	0.4762
	<b>industrie agroalimentaire</b>			0.4264	0.0775	0.4719	0.0862
	<b>industrie mécanique métallurgie</b>			-0.9235	0.0009	-0.5963	0.0592
	<b>réparation automobile</b>			0.5124	0.1518	0.6166	0.1369
	<b>services aux particuliers</b>			0.5763	0.0397	0.3629	0.2222
	<b>services aux entreprises</b>			-0.4450	0.3743	-0.5686	0.2659
	<b>autres</b>			-0.2954	0.2943	-0.2405	0.4194
<b>entreprise</b>	<b>association</b>			0.4532	0.3579	0.4491	0.3626
	<b>entreprise publique</b>			-0.4222	0.2632	-0.4070	0.2904
<b>privée</b>	<b>fonction publique</b>			-0.4749	0.2924	-0.8674	0.0762

	autres			0.2552	0.7934	0.4811	0.6189
établissement scolaire	ANPE					0.1832	0.4533
	annonce presse ou Internet					0.3562	0.0961
	agence d'intérim					-1.9595	<.0001
	Par relation					0.0621	0.6954
	autres					0.7320	0.0009
suite à un stage	en écrivant ou en se présentant					-0.1655	0.4632
	Par concours de fonction public					1.5730	0.0743
	suite à un 1er CDD					-0.2015	0.7199
	suite à un contrat d'apprentissage					0.3455	0.4087
	suite à un contrat intérim					0.6157	0.1355
	autres					0.2538	0.5216
test LR		8.1884	0.0423	132.9229	<.0001	254.3521	<.0001

**ANNEXE 3 :**

modèles économétriques		logit		modèle 4		Modèle 5	
Références	variables actives	coefficients	p-value	coefficients	p-value		
<b>Homme</b>	<b>femme</b>	7.5286	0.0061	-0.3018	0,0067		
<b>marié*au moins un enfant</b>	<b>célibataire*pas enfant</b>	-0.2013	0.0579	-0.1652	0.1549		
<b>Français</b>	<b>étrangère</b>	0.3514	0.2618	0,0882	0,8232		
	<b>constante</b>	0.8600	0.3537	-2.3995	0,8964		
<b>200 salariés Et Plus</b>	<b>aucun</b>	-1.0551	0.1557	-0.9047	0,2613		
	<b>moins de 10 salariés</b>	0.4076	0.0439	0,3829	0,0847		
	<b>entre 10 et 49 salariés</b>	0.4218	0.0392	0,3776	0,0891		
	<b>entre 50 et 99 salariés</b>	0.3688	0.2053	0,2121	0,5256		
	<b>entre 100 salariés et 199</b>	0.4601	0.1621	0,5128	0,1421		
	<b>inconnu</b>	-0.7859	0.0447	-0.6858	0,092		
<b>T R A N S P O R T</b>	<b>administration</b>	-0.4994	0.3894	-0.3769	0,5425		
	<b>agriculture</b>	-0.4712	0.1960	-0.4004	0,3676		
	<b>autres industries</b>	0.0120	0.9655	-0.0836	0,7868		
	<b>bâtiment du gros œuvre</b>	0.2203	0.4905	-0.0370	0,9203		
	<b>bâtiment du second degré</b>	0.4701	0.1218	0,3947	0,2466		
	<b>enseignement, santé</b>	-0.6099	0.2141	-0.6330	0,2504		
	<b>hôtellerie restauration</b>	0.7513	0.0020	0,7103	0,0092		
	<b>immobilier commerce</b>	-0.1733	0.5086	-0.2873	0,3125		
	<b>industrie agroalimentaire</b>	0.5119	0.0677	0,5175	0,1054		
	<b>industrie mécanique métallurgie</b>	-0.6307	0.0498	-0.5234	0,1327		
	<b>réparation automobile</b>	0.6536	0.1169	0,5816	0,1877		
	<b>services aux particuliers</b>	0.2858	0.3474	0,4465	0,1779		
	<b>services aux entreprises</b>	-0.6351	0.2234	-0.4885	0,3738		
	<b>autres</b>	-0.2471	0.4169	-0.4914	0,1725		
<b>Entreprise Privée</b>	<b>association</b>	0.4871	0.3349	0,1418	0,822		
	<b>entreprise publique</b>	-0.3969	0.3069	-0.4900	0,2804		
	<b>fonction publique</b>	-0.8764	0.0783	-0.7568	0,1909		
	<b>autres</b>	0.4843	0.6216	0,7365	0,4962		
<b>établissement Scolaire</b>	<b>ANPE</b>	0.1816	0.4624	0,0588	0,8311		
	<b>annonce presse ou Internet</b>	0.3915	0.0757	0,3775	0,12		
	<b>agence d'intérim</b>	-2.0060	<.0001	-1.8786	<.0001		
	<b>par relation</b>	0.0824	0.6078	0,1164	0,5095		
	<b>autres</b>	0.7485	0.0008	0,7674	0,002		

suite A  Un Stage	en écrivant ou en se présentant	-0.2059	0.3669	-0.3041	0,2203
	par concours de fonction public	1.6184	0.0663	1,4504	0,1058
	suite à un 1er CDD	-0.2239	0.6945	-0.2992	0,6415
	suite à un contrat d'apprentissage	0.3131	0.4641	0,3272	0,4725
	suite à un contrat intérim	-0.5944	0.1546	-0.8192	0,0759
	autres	0.2434	0.5473	0,5745	0,2579
père cadre	agent de maîtrise technicien	0.5420	0.1091	0,5212	0,1497
	agriculteur	-0.3285	0.5220	-0.3706	0,493
	artisan commerçant	0.3596	0.1562	0,2985	0,2836
	autre	-0.00451	0.9867	-0.2558	0,4074
	donnée inconnue	-0.4011	0.3268	0,0706	0,8845
	employé	-0.1584	0.3893	-0.1460	0,4691
	ouvrier	-0.1324	0.4769	-0.1954	0,346
mère cadre	agent de maîtrise technicien	0.4505	0.4208	0,4838	0,3981
	agriculteur	1.1383	0.1306	1,1846	0,1354
	artisan commerçant	-0.5669	0.2248	-0.6153	0,2186
	autre	-0.1488	0.5138	-0.0358	0,8883
	donnée inconnue	-0.2307	0.6254	-0.1839	0,7149
	employé	-0.1608	0.4267	-0.1881	0,3895
	ouvrier	-0.3669	0.2083	-0.3590	0,2577
CAP (dernier Diplôme Obtenu)	BAC Technologique			1,4157	0,9389
	BEP			1,8642	0,9194
	BTS			1,5681	0,9323
	DEUG			-6.4025	0,9566
	autres			1,025	0,9557
	BAC Général			-8.5428	0,959
	BAC Professionnel			1,6931	0,9268
	Brevet			1,4218	0,9385
	Brevet de technicien			1,3492	0,9417
	Brevet professionnel			2,0111	0,9131
	Mention complémentaire			0,5585	0,9758
test LR	264.1042	<.0001	236,2217	<.0001	