

## **Les déterminants du salaire net médian mensuel à la sortie d'un master universitaire en France en 2017**

### **Abstract**

Ce projet analyse les déterminants du salaire net médian mensuel des emplois à temps plein exprimé en euros, à la sortie d'un master universitaire en France en 2017. Pour cela, une étude économétrique est appliquée à partir de données provenant de l'opération nationale française de collecte de données sur l'insertion professionnelle des diplômés de Master français. Les résultats économétriques montrent un impact négatif significatif entre le salaire à la sortie d'un master et le pourcentage de femmes qui ont suivi ce master. Cependant, le taux d'insertion dans un master et le taux d'emplois cadres dans un master ont un impact significatif positif avec le salaire à la sortie d'étude. L'étude économétrique montre aussi l'importance du domaine d'étude et de la zone géographique où sont suivis les masters dans la détermination du salaire net médian. En effet, les résultats montrent que celui-ci est en moyenne plus élevé dans la zone géographique Île-de-France ainsi que dans le domaine droit, économie et gestion.

## **Introduction**

Avec un nombre d'inscriptions dans l'enseignement supérieur français multiplié par 8 environ entre 1960 et 2017, passant de 310 000 à 2 609 700, et comptabilisant 61% de ces inscriptions dans les universités de France métropolitaine et des départements d'outre-mer, les étudiants français sont de plus en plus nombreux à faire des longues études au fil du temps. Selon Lewandowski (2015), cela est dû au fait que le master reste la voie royale pour accéder à l'embauche, ce qui implique les étudiants à continuer leurs études jusqu'à celui-ci. De plus, Delès (2013) montre que le salaire médian augmente avec le niveau de diplôme. En parallèle, Gedalge, Hussenot et Huet (2023) affirment que les individus travaillent pour l'argent. Par conséquent, en liant ces trois informations, il est probable d'affirmer que les individus poursuivent leurs études jusqu'au master afin d'obtenir un premier salaire élevé.

Ainsi, l'étude des déterminants du salaire net médian mensuel à la sortie d'un master universitaire en France en 2017 vise à identifier les facteurs qui influent sur les niveaux de rémunération des diplômés de masters universitaires en France à leur entrée sur le marché du travail. Les objectifs principaux de cette étude sont donc d'identifier les facteurs qui influent sur le salaire post-master afin d'analyser les disparités salariales selon les différentes caractéristiques que proposent ces masters.

D'après la littérature, les emplois de cadres sont mieux rémunérés que les autres catégories socio-professionnelles (Charruau et Épaulard (2017)). La zone géographique et le domaine d'étude jouent aussi un rôle crucial dans la détermination du salaire post-master (Charruau et Épaulard (2017) ; Martinelli et Prost (2010)). En revanche, les académies selon Giret et Goudard (2010) ont un effet généralement faible sur le salaire à la sortie d'étude. Par conséquent, les académies ne jouent pas un rôle assez important pour souligner des disparités salariales post-master entre elles selon ces deux auteurs. De plus, selon Caro, Checcaglini et Guironnet (2022), les jeunes diplômés ont des salaires plus faibles lorsqu'ils sont dans des régions où le taux de chômage est plus élevé que la moyenne nationale. Un lien négatif relie donc le salaire à la sortie d'étude et le taux de chômage régional. Par ailleurs, le salaire post-master est lié négativement avec le genre féminin selon Kerr, Olivetti, Barth et Goldin (2017).

Ce projet étudie les déterminants du salaire net médian mensuel des emplois à temps plein à la sortie d'un master universitaire en France en 2017. Les résultats de l'étude économétrique montrent que le taux d'insertion dans un master, et le taux d'emplois cadres dans un master ont un impact significatif positif avec le salaire à la sortie d'étude. À contrario, celui-ci est impacté négativement de manière significative avec le taux de femmes. Quant au domaine et à la zone géographique, l'étude économétrique révèle leur importance dans la détermination du salaire post-master.

Ainsi, ce projet se divise en 4 parties. Tout d'abord, (1) une présentation de la base de données (2) accompagnée par la suite de sa méthode d'analyse sont présentées. Puis, (3) la revue de littérature de chaque variable est présentée avec son étude statistique ainsi que (4) son étude économétrique.

## **Chapitre 1 : Présentation de la base de données**

La base de données, créée par le département des outils d'aide à la décision, provient du site Open Data Enseignement supérieur et Recherche. Sa date de création est le 24 avril 2014. Elle est modifiée régulièrement. Elle regroupe les diplômés de master de 2010 à 2020 des universités de France métropolitaine et des DOM (hors Paris-Dauphine et l'Université Gustave Eiffel pour les diplômés 2020), de nationalité française, issus de la formation initiale, entrés immédiatement et durablement sur le marché de l'emploi après l'obtention de leur diplôme. L'objectif de ses créateurs est d'enquêter sur la situation professionnelle d'une part, 18 mois après obtention du diplôme et d'une autre part 30 mois après obtention du diplôme, des diplômés français de master français. Les enquêtes sont réalisées par les universités dans le cadre d'une charte dont les dispositions visent à garantir la comparabilité des résultats entre les établissements. Le traitement et la synthèse sont effectués au sein du ministère de l'Enseignement supérieur et de la Recherche (MESR) par la sous-direction des systèmes d'information et des études statistiques (SIES).

Initialement, la base de données comporte 19 604 observations pour 32 variables et traite 11 années (de 2010 à 2020). Après la sélection des variables retenues à l'aide de la littérature, elle est constituée de 811 observations pour 9 variables. La base de données porte uniquement sur l'année 2017. Une observation correspond à un master universitaire français. Une variable décrit une caractéristique de ces masters.

La problématique traite des différences salariales à la sortie d'un master universitaire en France. Pour l'analyser, 7 variables explicatives sont utilisées. D'après la littérature, celles-ci sont des déterminants potentiels pour expliquer la variable d'intérêt (Y) qui est le salaire net médian mensuel, exprimé en euros, des emplois à temps plein à la sortie d'un master universitaire en France, en 2017.

Le tableau 1 détaille les statistiques univariées de la variable d'intérêt Y. La moyenne est légèrement supérieure à sa médiane ce qui montre qu'il y a des salaires élevés qui tirent la moyenne vers le haut. Le 1er quartile et le 3ème quartile se trouvent à un écart semblable de la médiane (-200€ et +180€). Ceci prouve une certaine homogénéité de la variable Y. L'écart-type est de 260,64€. Il montre donc un certain étalement de la variable Y autour de la moyenne. Cependant, aucune valeur n'est aberrante.

Y	Minimum	1er quartile	Médiane	3ème quartile	Maximum	Moyenne	Écart-type
<b>Salaire net médian mensuel des emplois à temps plein à la sortie d'un master universitaire en France en 2017 (en €)</b>	1360	1750	1950	2130	3150	1960,69	260,64

*Tableau 1 : Statistiques univariées de la variable d'intérêt Y*

Les variables explicatives sont regroupées en 2 catégories.

### 1) Variables explicatives quantitatives :

- Le taux de femmes qui est le nombre de femmes diplômées sur le nombre total de diplômés au sein d'un master.
- Le taux d'insertion qui est le nombre de diplômés occupant un emploi, quel qu'il soit, sur le nombre total de diplômés présents sur le marché du travail au sein d'un master.
- Le taux de diplômés boursiers qui est le nombre de diplômés boursiers sur le nombre total de diplômés au sein d'un master
- Le taux de chômage régional qui est le nombre de chômeurs par régions sur le nombres total d'actifs par région
- Le taux d'emplois cadre qui est le nombre de diplômés occupant un emploi de cadre sur le nombre total de diplômés au sein d'un master.

### 2) Variables explicatives qualitatives

- Le domaine qui représente la branche d'étude des diplômés.

Les modalités de cette variable sont les suivantes : « droit, économie et gestion », « lettres, langues, arts », « masters enseignement », « sciences humaines et sociales », « sciences, technologies et santé ».

- L'académie qui représente le lieu d'étude des étudiants.

La variable académie était initialement composée de 24 académies qui ont été regroupées en 5 catégories en fonction de leur zone géographique. Le tableau 2 illustre ce changement.

Zones géographique	Académies
Nord-ouest	Normandie / Rennes / Nantes / Orléans-Tours / Poitiers
Nord-est	Lille / Amiens / Reims / Nancy-Metz / Strasbourg / Dijon / Besançon
Sud-ouest	Bordeaux / Limoges / Toulouse / Montpellier / Clermont-Ferrand
Sud-est	Lyon / Grenoble / Aix-Marseille / Nice
Île-de-France	Créteil / Versailles / Paris

*Tableau 2 : regroupement des académies en fonction de leur position géographique*

Le tableau 3 décrit statistiquement les variables explicatives quantitatives.

Xi 1	Minimum	1er quartile	Médiane	3ème quartile	Maximum	Moyenne	Écart-type
Taux de femmes (en %)	2	48	61	72	93	58,04	19,04
Taux d'insertion (en %)	69	89	93	96	100	91,81	5,18
Taux de diplômés boursiers (en %)	4	27	32	37	50	30,64	12,57
Taux de chômage régional (en %)	6,7	7,025	7,1	8	10	7,79	7,78
Taux d'emplois cadre (en %)	17	55	69	84	100	68,73	18,17

*Tableau 3 : Statistiques des variables explicatives quantitatives ( Xi 1)*

Le tableau 4 décrit statistiquement les variables explicatives qualitatives de notre base de données.

Xi 2	Modalités	Effectif	Fréquence (en %)
<b>Domaine</b>	Droit, économie et gestion	292	36,05%
	Lettres, langues, arts	38	4,69%
	Masters enseignement	63	7,78%
	Sciences humaines et sociales	168	20,74%
	Sciences, technologies et santé	249	30,74%
<b>Académie</b>	Nord-ouest	170	20,99%
	Nord-est	168	20,74%
	Sud-ouest	135	16,66%
	Sud-est	149	18,40%
	Île-de-France	188	23,21%

Tableau 4 : Statistiques des variables explicatives qualitatives (Xi 2)

## **Chapitre 2 : Méthode d'analyse de la base de données**

Cette partie consiste à expliquer les méthodes utilisées pour réaliser l'analyse statistique bivariée du lien entre le salaire net médian mensuel des emplois à temps pleins à la sortie d'un master universitaire en France en 2017 et les variables explicatives.

### **1) Cas d'une variable explicative quantitative :**

Pour tester la significativité du coefficient de corrélation, un test de Pearson est réalisé avec les hypothèses  $H_0$  : « Coefficient de corrélation nul » contre  $H_1$  : « Coefficient de corrélation non nul ». Pour tous les tests réalisés, est fixée une erreur de première espèce à 5%, qui désigne la probabilité de rejeter l'hypothèse nulle alors qu'elle est vraie. Ensuite, le calcul de la P-valeur, comparé à l'erreur de première espèce, permet de rejeter ou non l'hypothèse nulle. La conclusion consiste ainsi à dire s'il existe une corrélation ou non entre les deux variables quantitatives étudiées, avec un risque d'erreur de 5%.

### **2) Cas d'une variable explicative qualitative :**

Pour étudier la dépendance ou non d'une variable quantitative et qualitative, un test d'analyse de variance est utilisé (ANOVA à un facteur). On définit l'hypothèse nulle  $H_0$  : « égalité des moyennes » contre l'hypothèse non nulle  $H_1$  : « une moyenne s'écarte des autres ». L'erreur de

première espèce est toujours fixée à 5%. De même, la comparaison de la P-valeur avec l'erreur de première espèce, permet de conclure sur la corrélation entre les variables étudiées.

## **Chapitre 3 : Littérature économique et étude empirique**

*Pour simplifier l'écriture des résultats statistiques, le salaire net médian mensuel des emplois à temps plein exprimé en euros, à la sortie d'un master universitaire en France en 2017 équivaut pour toutes les parties « Étude Statistique » au salaire à la sortie d'étude.*

### **I) Le salaire et l'académie**

#### **1) Littérature**

Giret et Goudard (2010) réalisent une étude en France sur la mesure de l'effet de l'établissement sur le salaire des diplômés du supérieur. Cette étude révèle que l'effet propre de l'établissement est généralement faible, car la détermination des salaires est expliquée par des variables individuelles. Par conséquent, une académie, englobant un certain nombre d'établissements, ne parvient pas à expliquer les salaires.

Cependant, le lieu géographique joue un rôle dans la détermination des salaires (Charruau et Épaulard). Le découpage de ces zones géographique est fait par rapport aux régions Françaises. Chaque académie appartient à une zone géographique. Plusieurs effets déterminants sont distingués. Tout d'abord, dans une zone avec un marché du travail suffisamment grand, la diversification de l'offre et de la demande d'emplois permet un meilleur appariement. De plus, la concentration spatiale des firmes favorise la circulation des connaissances entre elles, ce qui encourage l'innovation, souvent appelée effet « Silicon Valley ». La concentration spatiale des firmes augmente également l'offre de sous-traitants dans la zone concernée, avec une présence accrue de comptables, entreprises de transports ou de nettoyage, permettant aux entreprises d'accéder à des services nécessaires à moindre coût. Ces effets favorisent la performance globale de l'économie, entraînant des salaires plus élevés dans ces zones géographiques dynamiques. Des salaires plus élevés peuvent également s'expliquer par le fait que seules les firmes et les employés les plus productifs y demeurent, tandis que les autres s'éloignent vers des zones où la concurrence est moins rude.

De plus, dans leur étude en France sur « L'empreinte des territoires sur le salaire des jeunes débutants », Caro, Checcaglini et Guironnet (2022) démontrent que résider dans une région affichant un salaire médian supérieur confère un avantage en termes de rémunération pour les jeunes diplômés. Cet avantage est accentué à Paris. Ces zones d'emplois témoignent d'un dynamisme marqué en matière de démographie et de marché du travail. Dans ces régions, les emplois se spécialisent dans certains domaines, entraînant une production accrue de richesse et, par conséquent, une rémunération plus substantielle pour ceux qui y travaillent.

Malgré le constat de Giret et Goudard (2010) selon lequel l'établissement ne joue pas directement un rôle dans la détermination des salaires, Charruau et Épaulard (2017) démontrent que la zone géographique, donc l'académie, joue un rôle important tout comme Caro, Checcaglini et Guironnet (2022).

## **2) Étude Statistique**

Les académies sont divisées en 5 zones géographiques. Les moyennes par zone géographique du salaire à la sortie d'étude sont : Île-de-France (2148€), Nord-est (1954€), Nord-ouest (1869€), Sud-est (1913€), Sud-ouest (1876€). L'Île-de-France possède donc, en moyenne, un salaire à la sortie d'étude plus élevé comparé aux autres zones géographiques. Ces zones géographiques ont quant à elles, en moyenne, un salaire à la sortie d'étude semblable entre elles.

Afin d'étudier si des disparités entre les zones existent, un test ANOVA est réalisé. Les conditions de validités sont vérifiées puisque la taille de l'échantillon est supérieure à 30. Le test asymptotique est donc réalisable. La P-valeur du test d'égalité des variances est égale à  $9,675e-07$  qui est inférieur à 5%, donc l'homogénéité des variances n'est pas vérifiée au niveau de risque 5%. La P-valeur du test ANOVA est de  $2,2e-16$  ce qui est inférieur à 5%, donc l'hypothèse d'égalité est rejetée à un niveau de confiance de 95%. Cela signifie qu'au moins deux modalités ont des moyennes différentes.

Afin de trouver quelles modalités ont des variances différentes, les tests de comparaison multiples des moyennes est réalisé sur les variables 2 à 2. On fait le test  $H_0^{ij}: M_i=M_j$  contre  $H_1^{ij}: M_i \neq M_j$ . Les P-valeurs sont inférieures à 5% entre la zone Île-de-France et toutes les autres zones géographiques, entre nord-est et nord-ouest ainsi qu'entre nord-est et sud-ouest. Cela signifie donc des différences significatives de moyennes entre ces zones.

Ainsi, la zone géographique a un impact sur le salaire à la sortie d'étude en France. Ces résultats correspondent avec ceux de Charruau et Épaulard (2017) ainsi que Caro, Checcaglini et Guironnet (2022).

## **II) Le salaire et le taux de chômage**

### **1) Littérature**

Aaronson et Sullivan (2000) s'appuient sur des données américaines et soulignent une corrélation négative entre le taux de croissance du salaire et le taux de chômage. Plus précisément, ils mettent en exergue que les États avec des taux de chômage élevés présentent des salaires plus faibles que les autres.



Caro, Checcaglini et Guironnet (2022), montrent que les jeunes diplômés, en France, ont des salaires plus faibles lorsqu'ils sont dans des régions où le taux de chômage est plus élevé que la moyenne nationale.

Néanmoins, Nikulin (2015) montre qu'il n'y a pas de corrélation entre le salaire et le taux de chômage dans les pays européens (exception pour la Hongrie).

## **2) Étude statistique**

Afin d'observer la relation entre le salaire à la sortie d'étude et le taux de chômage, un calcul du coefficient de corrélation entre ces deux variables est effectué. Celui-ci est égal -0,10.

Le test de Pearson est utilisé afin de vérifier la significativité des résultats obtenus précédemment. La P-valeur est égale à 0,006718 ce qui est inférieur à 0,05 donc on rejette l'hypothèse  $H_0$  au niveau 5%. Cela signifie qu'il existe une relation linéaire négative significative entre le salaire à la sortie d'étude et le taux de chômage. Ce résultat est en accord avec la littérature.

## **III) Le salaire et le taux d'emploi cadre**

### **1) Littérature**

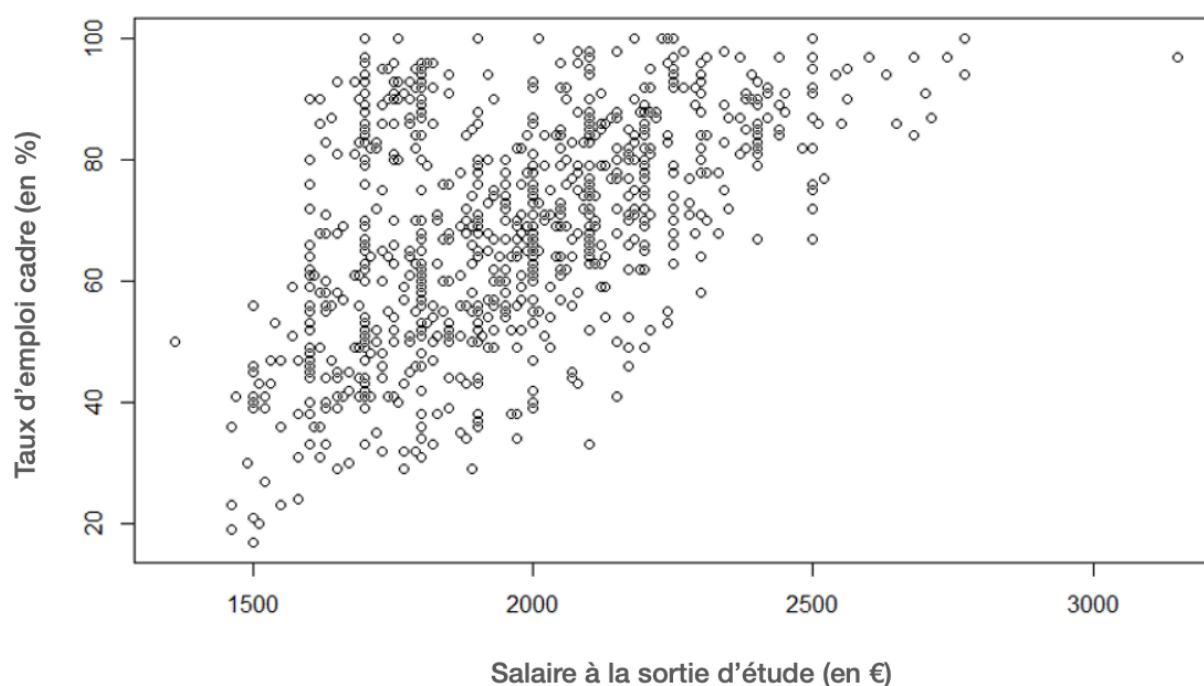
Le fait d'être cadre se révèle être un élément déterminant du niveau salarial d'un individu en France (Charruau et Épaulard, 2017). En Île-de-France, les hommes et les femmes cadres sont les mieux rémunérés. Aeberhardt et Pouget (2006) partagent le même avis. Dans un article pour l'INSEE, ils expliquent que le salaire est étroitement lié à la catégorie socioprofessionnelle. Un homme cadre gagne 50% de plus qu'un ouvrier ou un employé. Lhéritier (1992) explique que le fait que les cadres aient un salaire plus élevé est dû à un degré plus fort de capital humain. Selon Becker (1993), le capital humain se définit par « un actif, un patrimoine, une réserve de connaissances et d'expériences accumulées tout au long d'une vie grâce à des investissements visant à accroître sa productivité future et, par conséquent, son salaire ». En effet, les cadres ont généralement accumulé plus de capital humain car ils sont plus diplômés. Le lien peut être établi avec la formation suivie par l'étudiant. Un master qui forme davantage de futurs cadres (taux d'emploi cadre) est lié à un salaire plus élevé.

Battagliola (2001) note toutefois que les femmes parviennent difficilement aux positions dirigeantes. A ce titre, elles ne représentent qu'un tiers des emplois cadres en France. Cependant elles ne sont que 3 % des numéros « un » (PDG) et 6 % des numéros deux (DG) dans les 2 000 entreprises de plus de 500 salariés. Ce faisant ainsi, le taux de femme semble liée au taux de cadre, lui-même lié au salaire. Plus précisément, il se pourrait qu'un taux de femmes élevé dans un master conduise à un salaire médian plus faible que celui d'un autre master avec un taux moins élevé.

## 2) Étude statistique

Afin d'observer la relation entre le salaire à la sortie d'étude et le taux d'emploi cadre, un calcul du coefficient de corrélation entre ces deux variables est effectué. Celui-ci est égal à 0,52.

Pour vérifier sa significativité, un test de Pearson est réalisé. La P-valeur est égale à  $2,2e-16$ , ce qui est inférieur à 0,05. On rejette donc l'hypothèse  $H_0$  au niveau 5%. De ce fait, il y a une relation linéaire positive significative entre le salaire à la sortie d'étude et le taux d'emploi cadre. Ce résultat est en accord avec la littérature qui trouve un lien positif entre le taux d'emploi cadre et le salaire médian. De plus, la relation linéaire positive est visible sur le nuage de points 2. En effet, une droite ascendante de régression peut être tracée. Un point aberrant apparaît sur ce nuage de points. Il correspond à un master d'Île-de-France dans le domaine d'étude droit, économie et gestion.



*Nuage de points 2 : le salaire à la sortie d'étude (en €) en fonction du pourcentage d'emploi cadre.*

## IV) Le salaire et le taux d'insertion

### 1) Littérature

Selon Gaini, Leduc, Vicard (2016), il n'existe pas de corrélation entre le taux d'insertion et le salaire. En effet, les diplômés sont généralement rémunérés au salaire minimum. Lequel est indépendant du taux d'insertion.

## **2) Étude statistique**

Afin d'observer la relation entre le salaire à la sortie d'étude et le taux d'insertion, un calcul du coefficient de corrélation est effectué entre ces deux variables. Celui-ci est égal à 0,41.

Le test de Pearson permet de vérifier la significativité du résultat précédent. La P-valeur du test est égale à  $2,2e-16$  ce qui est inférieur à 0,05. On rejette donc l'hypothèse  $H_0$  au niveau 5%. Ainsi, il existe une forte relation linéaire positive significative entre le salaire à la sortie d'étude et le taux d'insertion.

Ces résultats ne correspondent pas à ceux de la revue de littérature. Une explication possible à cette différence est que l'étude réalisée par Gaini, Leduc, Vicard (2016) a été réalisée sur des données datant de 1982 à 2010 alors que les données utilisées pour cette étude datent de 2017. Ainsi, le résultat trouvé par Gaini, Leduc, Vicard (2016) peut donc avoir changé au cours du temps.

## **V) Le salaire et le domaine d'étude**

### **1) Littérature**

Le domaine d'études des étudiants semble jouer un rôle dans le salaire obtenu en sortie d'études. Martinelli et Prost (2010) montrent que le domaine d'étude joue un rôle majeur dans la détermination des salaires en France. Les diplômés d'un master en physique, mathématiques, droit, sciences politiques, ou encore commerce/vente ont un salaire médian d'environ 2000 euros, tandis que les diplômés d'un master en histoire, géographie, sociologie, psychologie, ou encore arts gagnent autour de 1500 euros en début de carrière.

Aeberhardt et Pouget (2006) affirment que ces disparités salariales s'expliquent par des rendements différents du travail causés par la différence des technologies utilisées, propres à chaque secteur. Des caractéristiques institutionnelles et conjoncturelles peuvent aussi varier d'un domaine à l'autre et impacter les salaires que perçoivent les jeunes diplômés. Dans certains domaines, notamment le domaine financier, les entreprises sont capables de capter une importante partie du surplus économique et ainsi d'en distribuer une partie à leurs salariés, ce qui augmente leurs salaires (Lhéritier, 1992). Il explique que cela est dû au fait que ces entreprises sont souvent en situation de quasi-monopole ou d'oligopole.

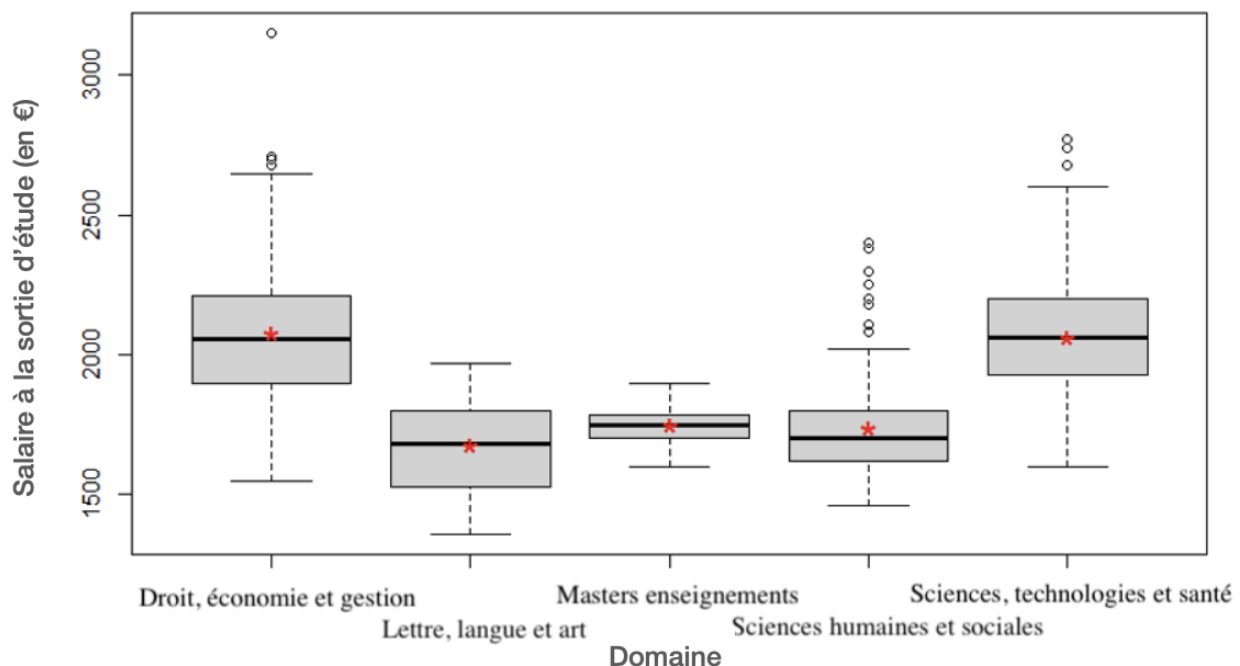
### **2) Étude statistique**

Les domaines d'étude sont divisés en 5 modalités. Les moyennes par modalité du salaire à la sortie d'étude sont : Droit, économie et gestion (2080€), Lettres, langues, arts (1681€), Masters enseignement (1750€), Sciences humaines et sociales (1740€), Sciences, technologies et santé (2067€). Les moyennes sont donc différentes entre ces modalités.

Afin d'étudier la liaison entre les groupes, un test ANOVA est effectué. Les conditions de validité sont vérifiées puisque l'échantillon est supérieur à 30. Le test asymptotique est donc réalisable. La P-valeur du test d'égalité des variances est de 6,814e-15 ce qui est inférieur à 5%. Par conséquent, l'homogénéité des variances n'est pas vérifiée au niveau 5%. La P-valeur du test ANOVA est de 2,2e-16 ce qui est inférieur à 5%. L'hypothèse d'égalité est donc rejetée à 95% de confiance. De ce fait, certaines des modalités ont une moyenne différente des autres.

Afin de trouver quelles modalités ont des variances différentes, les tests de comparaison multiples des médianes est réalisé sur les variables 2 à 2. On fait le test  $H_0^{ij}: M_i=M_j$  contre  $H_1^{ij}: M_i \neq M_j$ . Les P-valeurs sont toutes inférieures à 5% sauf celles entre « Droit, économie et gestion » et « Sciences, technologie et santé » ainsi qu'entre « Master enseignements » et « Sciences Humaines et sociales ». Ainsi, le domaine d'étude influe sur le salaire à la sortie d'étude.

Ce résultat correspond avec la littérature. De plus, sur la boîte à moustache 1, ce résultat est visible car le domaine « Droit, économie et gestion » et le domaine « Sciences, technologie et santé » montrent un salaire moyen plus élevé comparé aux autres domaines ainsi qu'un salaire moyen semblable entre eux.



Boite à moustache 1 : Le salaire à la sortie d'étude (en €) en fonction du domaine d'étude.

## **VI) Le salaire et le taux de femmes**

### **1) Littérature**

Kerr, Olivetti, Barth et Goldin (2017) mettent en évidence les inégalités salariales liées au genres, et donc une corrélation entre le sexe d'un individu et son salaire. Plus particulièrement, ils montrent que ces inégalités sont observées entre des hommes et des femmes exerçant une même profession, et surviennent en grande partie à la suite de la naissance du premier enfant.

En utilisant des données britanniques, cette idée est soutenue par Brown, Roberts et Taylor (2011) qui montrent que la charge d'un enfant explique en très grande partie les inégalités de salaires entre hommes et femmes.

Loprest (1992) montre que le changement de travail, plus fréquent chez les femmes, est une explication de l'infériorité des salaires que perçoivent les femmes. En effet, cette instabilité limite les possibilités d'ascension hiérarchique au sein des entreprises.

Kerr, Olivetti, Barth et Goldin (2017) soulignent aussi que les hommes sont davantage employés dans les postes de travail à hauts revenus, et que des ascensions dans la hiérarchie de l'entreprise sont bien plus fréquentes chez les hommes que chez les femmes. D'autant plus que les théories dites du « plafond de verre » montrent que l'ascension hiérarchique de la femme est possible mais toutefois limitée à un certain stade au niveau hiérarchique.

Le plafond de verre peut se définir comme l'ensemble des obstacles visibles et invisibles qui séparent les femmes du sommet des hiérarchies professionnelles et organisationnelles. Laufer (2005) explicite le concept de plafond de verre en s'appuyant sur une étude réalisée en France, et donne un certain nombre de déterminants de ce dernier. Pour ce faire, elle s'appuie sur une approche plutôt sociologique, en traitant par exemple du rôle des normes, de la culture ou encore de la famille dans ce que l'on pourrait qualifier d'une sorte de déterminisme lié au plafond de verre. De ce fait, les femmes éprouvent des difficultés à accéder à des postes à responsabilités, à hauts revenus, ce qui borne en quelque sorte leur potentiel revenu relativement à celui des hommes.

En s'appuyant sur des données américaines, Blau et Kahn (2017) confirment et précisent que la position dans la hiérarchie de l'entreprise ainsi que l'industrie sont des déterminants clés de l'écart de salaires entre hommes et femmes.

Battagliola (2001), montre en s'appuyant sur des données françaises, que la main d'œuvre féminine est différenciée de la masculine. Plus particulièrement, en raison de la naissance d'enfant et du congé parental accordé qui conduit à une période d'indisponibilité. Plus tard, c'est la gestion de la vie familiale qui porte préjudice à la femme. Finalement, ces raisons font que pour un certain nombre de postes et domaines d'activité, les femmes sont sous-représentées. A l'inverse, les femmes sont surreprésentées dans le secteur tertiaire, où les emplois à temps partiel sont prédominants. Ainsi, il est possible que le taux de femmes dans un master ait un effet sur le salaire médian perçu à

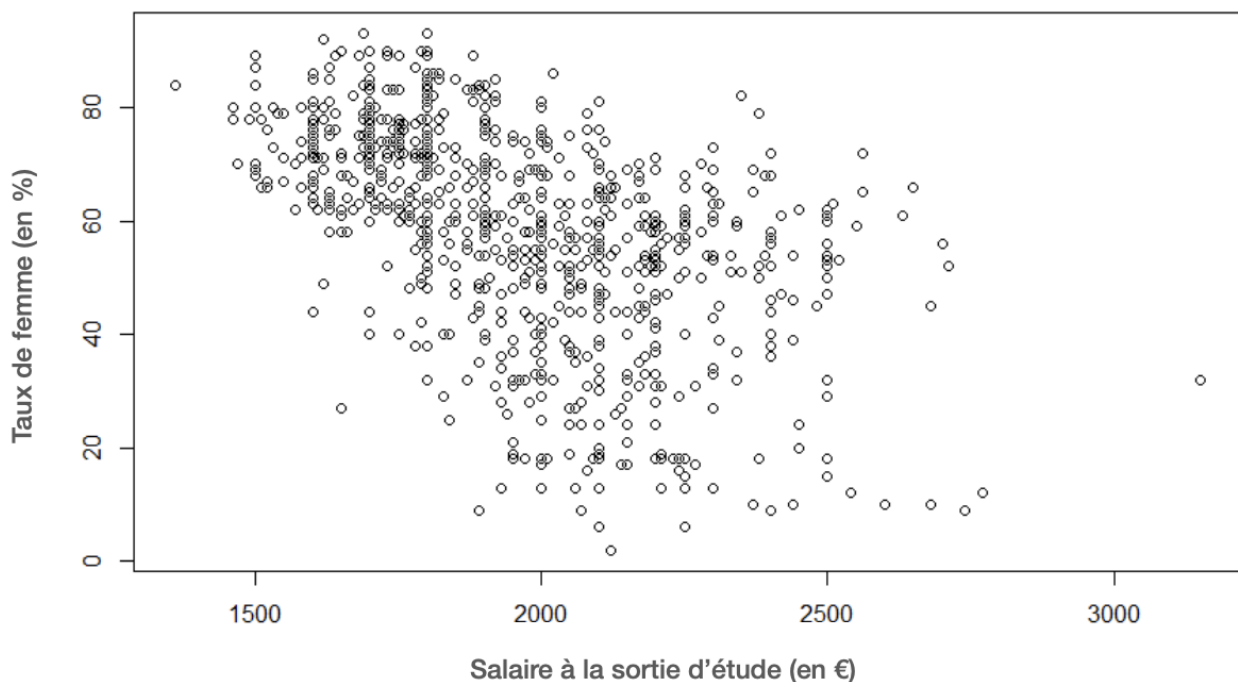
la sortie de ce dernier. En effet, puisque le taux de femmes est plus élevé dans certains domaines, et qu'entre les domaines les salaires ne sont pas les mêmes, le taux de femmes peut donc influencer le salaire à la sortie d'étude.

## 2) Étude statistique

Afin d'observer la relation entre le salaire et le pourcentage des femmes, un calcul du coefficient de corrélation entre ces deux variables est effectué. Celui-ci est égal à  $-0,53$ .

De manière à vérifier la significativité de ces résultats, un test de Pearson est réalisé. La P-valeur est égale à  $2,2e-16$  ce qui est inférieur à  $0,05$  donc l'hypothèse  $H_0$  est rejetée au niveau  $5\%$ . Cela veut dire qu'il existe une forte relation linéaire négative significative entre le salaire à la sortie d'étude et le taux de femmes.

De ce fait, plus le pourcentage de femmes dans une formation augmente, plus le salaire en sortie d'étude de cette formation est faible. Ce résultat coïncide donc avec la littérature. De plus, sur le nuage de points 2, la liaison linéaire négative est visible. Effectivement, une droite de régression descendante peut être tracée sur ce graphique. Cela signifie qu'il existe une relation linéaire négative significative entre le salaire à la sortie d'étude et le pourcentage des femmes. Ce résultat est en accord avec la littérature.



*Nuage de points 2 : le salaire à la sortie d'étude (en €) en fonction du pourcentage de femmes présentes dans la formation associée.*

## **VII) Le salaire et le taux de diplômés boursiers**

### **1) Littérature**

Selon Demeuse, Friant, Malaise (2014) les individus susceptibles de présenter des risques d'échec ou d'abandon plus importants à l'université sont notamment les étudiants boursiers et les étudiants issues de lycée dont l'indice socio-économique est faible. De ce fait, si un établissement est majoritairement fréquenté par des étudiants boursiers, qui sont parmi les étudiants avec le plus fort taux d'échec, les offreurs de travail peuvent considérer que le niveau de cette université est faible, et donc aussi les compétences acquises dans ses formations. Toutefois, Bourdon, Giret et Goudard (2011) montrent que le taux de boursier d'une université n'a pas d'effets négatifs sur le salaire de ses diplômés. Ils attestent seulement d'une corrélation négative du taux de boursier avec l'insertion des diplômés sur le marché du travail.

Nimer (2016), en s'appuyant sur des données libanaises, montre que les boursiers et non boursiers s'orientent vers des filières différentes. Dans son étude, elle observe que les non boursiers choisissent majoritairement de suivre des études dans la gestion, un domaine dans lequel les salaires sont élevés. De ce fait, le taux de boursier semble influencer sur le salaire médian, par le biais du domaine d'étude.

### **2) Étude statistique**

Afin d'observer la relation entre le salaire à la sortie d'étude et le taux de diplômés boursiers, un calcul du coefficient de corrélation entre ces deux variables est effectué. Celui-ci est égal à -0,30.

Le test de Pearson permet de vérifier la significativité de cette corrélation. La P-valeur est égale à  $2,2e-16$  ce qui est inférieur à 0,05 donc l'hypothèse  $H_0$  est rejetée au niveau 5%. Ainsi, il existe une relation linéaire négative significative entre le salaire à la sortie d'étude et le taux de diplômés boursiers.

Par conséquent, plus le taux de diplômés boursiers est élevé dans un master, plus les salaires en sortie de ce master sont faibles. Ce résultat n'est pas en accord avec la littérature.

## **Chapitre 4 : Étude économétrique**

*Pour simplifier l'écriture de l'étude économétrique, le salaire net médian mensuel des emplois à temps plein exprimé en euros, à la sortie d'un master universitaire en France en 2017 équivaut pour toutes la partie « Étude économétrique » au salaire à la sortie d'étude.*

## I) Présentation des variables

### 1) Présentation de la variable d'intérêt

La variable d'intérêt est le salaire net médian mensuel, exprimé en euros, des emplois à temps plein à la sortie d'un master universitaire en France, en 2017. Afin de l'expliquer, 7 variables explicatives sont utilisées.

### 2) Présentation des variables explicatives

Les variables explicatives quantitatives utilisées dans l'étude économétrique sont: le taux de chômage (« Tchom »), le taux d'emploi cadre (« Prcadre »), le taux d'insertion (« Tinsert »), le taux de diplômés boursier (« Prdiplom ») et le taux de femmes (« Prfemme »).

Les variables explicatives qualitatives utilisées dans l'étude économétrique sont l'académie, qui est regroupée en zone géographique et le domaine d'étude. La variable qualitative académie a 5 modalités et est codée en 4 variables indicatrices en prenant la zone géographique « Île-de-France » en modalité de référence. Quant au domaine, il existe aussi 5 modalités. Cette variable est codée en 4 variables indicatrices. La modalité de référence est « Droit, économie et gestion ». Le tableau 5 résume le codage de ces 2 variables qualitatives en variables binaires.

Variables	Modalités	Modalité de référence
<u>L'académie est déclinée en plusieurs modalités :</u> Île-de-France ("idf") Nord-Ouest ("no") Nord-Est ("ne") Sud-Ouest ("so") Sud-Est ("se")	"0" si le master n'appartient pas à la zone géographique  "1" si le master appartient à la zone géographique	Île-de-France
<u>Le domaine d'étude est décliné en plusieurs modalités:</u> Droit, Économie et Gestion ("deg") Lettre, langues et arts ("lla") Masters enseignement ("me") Sciences humaines et sociales ("shs") Sciences, technologies et santé ("sts")	"0" si le master n'appartient pas au domaine d'étude  "1" si le master appartient au domaine d'étude	Droit, Économie et Gestion

Tableau 5: Codage des variables qualitatives en variables binaires



### 3) Résultats attendus d'après la littérature

La revue de littérature effectuée dans le chapitre 3 donne des résultats sur les liens attendus entre les variables explicatives et la variable à expliquer.

Pour les variables qualitatives, celle-ci montre que l'académie et le domaine d'étude jouent un rôle important dans la détermination du salaire à la sortie d'étude. Effectivement, le salaire à la sortie d'étude est supérieur en Île-de-France et des disparités salariales existent en fonction du domaine étudié selon elle.

Quant aux variables quantitatives, le tableau 6 décrit les liens attendus par la revue de littérature et celles-ci.

Variables explicatives	Lien avec la littérature
Taux de femmes (en %)	Lien négatif avec le salaire à la sortie d'étude
Taux d'insertion (en %)	Aucun lien avec le salaire à la sortie d'étude
Taux de diplômés boursiers (en %)	Lien négatif avec le salaire à la sortie d'étude
Taux de chômage régional (en %)	Lien négatif avec le salaire à la sortie d'étude
Taux d'emplois cadre (en %)	Lien positif avec le salaire à la sortie d'étude

*Tableau 6 : Résultat attendu par la littérature pour les variables quantitatives*

## II) Étude des modèles de régression linéaire multiple

### 1) Présentation des modèles utilisés

Pour l'étude économétrique, le modèle utilisé est un modèle de régression linéaire multiple du type :  $y = X\beta + \epsilon$ . Il permet de modéliser la relation entre une variable dépendante  $y$  : le salaire net médian mensuel des emplois à temps plein exprimé en euros, à la sortie d'un master universitaire en France en 2017 et des variables explicatives. La méthode de régression linéaire multiple est appliquée à la base de données présentée dans le chapitre 1 afin d'estimer les coefficients de régression.

Pour cela, la méthode des moindres carrés ordinaire (MCO) est utilisée. Cette méthode consiste à minimiser la somme des carrés des écarts entre les valeurs prédites par le modèle et les valeurs observées. Les hypothèses statistiques sont vérifiées : les observations de la base de données

sont indépendantes et identiquement distribuées, les coefficients de régressions contenus dans le vecteur sont des nombres réels, les termes d'erreur sont indépendants entre eux et la variance des observations est strictement positive. De plus, on suppose vérifiée l'hypothèse d'homoscédasticité du terme d'erreur  $\epsilon$ . Ainsi, les coefficients de régression obtenus, contenus dans le vecteur  $\beta$ , sont sans biais et, selon le théorème de Gauss-Markov, l'estimateur par MCO est le meilleur estimateur linéaire sans biais (BLUE).

Pour pouvoir étudier sous plusieurs angles l'impact qu'ont les variables explicatives sur la variable dépendante, trois modèles de régression linéaire multiple sont utilisés. Deux modèles de régression linéaire multiple : un modèle Log-Level et un modèle avec des effets croisés entre certaines variables. De plus, la base de données est ensuite divisée en 2 échantillons : un échantillon contenant les observations avec un salaire supérieur au salaire médian et un échantillon contenant les observations avec un salaire inférieur ou égal au salaire médian. Une régression linéaire multiple Log-Level est appliquée à ces deux échantillons afin de comparer les coefficients de régression.

Le logiciel SAS Studio est utilisé dans l'étude économétrique. Celui-ci permet de donner les résultats présents dans le tableau 7. Ce tableau donne les résultats des coefficients de régressions et des erreurs standard pour chaque variable explicative dans chacun des modèles utilisés. Dans ce tableau, si une variable est significative à 1% alors elle est suivie de 3 étoiles (\*\*\*), 5% (\*\*), 10% (\*). Les erreurs standard sont entre parenthèses. Les effets croisés sont notés de la manière suivante : variable1\_variable2.

	Modèle Log-Level	Modèle Croisé	Modèle « Higher »	Modèle « Lower »
Intercept	7,29865*** (0,05710)	7,14817*** (0,06619)	7,31493*** (0,08285)	7,23384*** (0,05662)
Prfemme	-0,00180*** (0,00018023)	-0,00137** (0,00062812)	-0,00081628*** (0,00023679)	-0,00115*** (0,00021071)
Tinsert	0,00370*** (0,00053224)	0,00448*** (0,00052149)	0,00316*** (0,00078183)	0,00316*** (0,00051630)
Prdiplom	-0,00036941 (0,00042442)	0,00033532 (0,00042541)	0,00004678 (0,00056276)	0,00016226 (0,00043593)
Tchom	0,00007694 (0,00248)	0,00116 (0,00239)	-0,00008109 (0,00344)	0,00504** (0,00240)
Prcadre	0,00271*** (0,00017350)	0,00486*** (0,00050456)	0,00245*** (0,00029386)	0,00148*** (0,00016831)
Ne	-0,06827*** (0,00882)	-0,06725*** (0,00846)	-0,05032*** (0,01175)	-0,05958*** (0,00915)
No	-0,09903*** (0,00869)	-0,10372*** (0,00836)	-0,08577*** (0,01177)	-0,06684*** (0,00893)
Se	-0,07193*** (0,00818)	-0,07374*** (0,00787)	-0,06428*** (0,01085)	-0,04925*** (0,00849)

So	-0,09976*** (0,0098)	-0,10624*** (0,00943)	-0,08248*** (0,01305)	-0,07659*** (0,00991)
Shs	-0,13330*** (0,00672)	-0,07229* (0,04333)	-0,074*** (0,01402)	-0,08056*** (0,00692)
Sts	-0,07225*** (0,00673)	-0,24651*** (0,02728)	-0,05805*** (0,00826)	-0,02409*** (0,00802)
Lla	-0,10063*** (0,01164)	0,06668 (0,12258)	0,03440 (0,05547)	-0,06237*** (0,01005)
Me	-0,22611*** (0,01177)	-0,25142** (0,09965)	0	-0,11252*** (0,01170)
Prfemme_shs		-0,00050746 (0,00063016)		
Prfemme_sts		0,00327*** (0,00048002)		
Prfemme_lla		-0,00175 (0,00157)		
Prfemme_me		0,00078065 (0,00136)		
Prfemme_prcadre		-0,00002991*** (0,00000765)		
Prfemme_prcadre		-0,00004850*** (0,00000853)		

Tableau 7 : Résultats des modèles de régression linéaire multiple

## 2) Modèle de régression linéaire Log-Level

$$\text{Log}(\text{salnetmedian}) = \beta_0 + \beta_1 \text{prfemme} + \beta_2 \text{tinsert} + \beta_3 \text{prdiplom} + \beta_4 \text{tchom} + \beta_5 \text{prcadre} + \beta_6 \text{ne} + \beta_7 \text{no} + \beta_8 \text{se} + \beta_9 \text{so} + \beta_{10} \text{shs} + \beta_{11} \text{sts} + \beta_{12} \text{lla} + \beta_{13} \text{me} + \varepsilon$$

### a) Présentation du modèle

Tout d'abord, un modèle Log-Level est utilisé car étant donné que toutes les variables de la base de données sont des taux, cela permet d'exprimer la variation du salaire net médian en pourcentage.

La significativité générale du modèle doit être testée. Pour cela, tous les coefficients de régression sont testés simultanément. Les hypothèses sont : H0 : « Tous les coefficients de régression sont nuls » et H1 : « Au moins un des coefficients de régression est non nul ». La statistique de test suit une loi de Fisher, la p-valeur est inférieure à 0,0001 alors l'hypothèse H0 est

rejetée au niveau 1%. Ainsi au moins un des coefficients de régression est non nul, donc le modèle a au moins une variable explicative.

Le coefficient de détermination noté :  $R^2$  est un indicateur compris entre 0 et 1 permettant de mesurer la qualité d'ajustement d'un modèle. Il est calculé en divisant la somme des carrés expliqués par la somme des carrés totaux des observations. Cependant, le  $R^2$  n'est pas suffisant. Effectivement, cet indicateur ne peut qu'augmenter, même à la suite de l'ajout d'une variable non significative. Le  $R^2$ -ajusté est un indicateur similaire au  $R^2$  mais il applique une pénalité lorsqu'une variable non significative est ajoutée. Le calcul du  $R^2$ -ajusté est réalisé afin d'évaluer la qualité du modèle correctement.

Le  $R^2$  de ce modèle est égale à 0,7863 ce qui signifie que 78,63% de la variance de la variable dépendante Y est expliquée par les variables explicatives du modèle. Quant à lui, le  $R^2$ -ajusté est égale à 0,7829, donc 78,29% de la variance de la variable dépendante Y est expliquée par les variables significatives du modèle.

## **b) Résultats et interprétations**

La significativité de chaque variable est testée à l'aide de la statistique de Student et la p-valeur. Les hypothèses pour chaque test sont  $H_0$  : « le coefficient de régression est nul » et  $H_1$  : « le coefficient de régression est non nul ».

La p-valeur de la variable « Taux de femme » est inférieure à 0,0001. Alors, l'hypothèse  $H_0$  est rejetée au niveau 1% ce qui signifie que la variable est significative. Le taux de femme dans un master en France en 2017 a donc un impact négatif sur le salaire en sortie d'étude. Le coefficient de régression de la variable est égal à -0,0018. Donc, selon le modèle économétrique, pour un point de pourcentage supplémentaire de femme dans un master en France en 2017, le salaire en sortie d'étude de ce master diminue de 0,18%, toutes choses égales par ailleurs. Ce résultat correspond à ceux de Kerr, Olivetti, Barth et Goldin (2017) qui mettent en évidence une disparité salariale entre les hommes et les femmes.

La variable « Taux d'insertion » a une p-valeur inférieure à 0,0001 donc l'hypothèse  $H_0$  est rejetée au niveau 1%. Ainsi, la variable est significative pour le modèle, le taux d'insertion à la sortie du master en France en 2017 impacte donc le salaire en sortie d'étude. Le coefficient de régression est égal à 0,0037. De ce fait, selon le modèle économétrique, pour un point de pourcentage supplémentaire du taux d'insertion dans un master, le salaire à la sortie d'étude de ce master augmente de 0,37%, toutes choses égales par ailleurs. Ce résultat ne correspond pas avec la littérature. En effet, selon Gaini, Leduc, Vicard (2016), il n'existe pas de corrélation entre le taux d'insertion et le salaire. Cela est lié au fait que généralement les diplômés sont rémunérés au salaire minimum, peu importe le taux d'insertion. Une hypothèse de ce résultat peut être que certains masters notamment en Île-de-France avec certains domaines, comme principalement le domaine Droit, économie et gestion offrent des salaires supérieurs au salaire minimum et ont des taux

d'insertion élevés. Sachant que l'impact du taux d'insertion sur le salaire en sortie d'étude est relativement faible, il est alors possible que ces quelques observations ayant un salaire supérieur au salaire minimum rendent la variable « Taux d'insertion » significative dans ce modèle. Il serait alors intéressant de faire 2 régressions : une sur la partie basse des salaires en sortie d'étude et une sur la partie haute. Cela permettrait de se rendre compte de l'impact réel du taux d'insertion sur les données.

Pour la variable « Taux de diplômés boursiers », la p-valeur est égale à 0,3844 donc l'hypothèse  $H_0$  n'est pas rejetée. Le taux de diplômés boursiers dans un master en France en 2017 n'a donc pas d'impact sur le salaire en sortie d'étude. Ce résultat ne correspond pas avec ceux de Demeuse, Friant, Malaise (2014) qui montrent que les entreprises offrant des salaires supérieurs au salaire minimum peuvent se montrer réticentes à employer des diplômés d'universités présentant un fort taux de diplômés boursier supposant que les élèves ont des compétences moins intéressantes. Cependant, il correspond avec celui de Bourdon, Giret et Goudard (2011) qui montrent qu'il n'y a pas d'effet négatif du taux de diplômés boursier d'un master sur le salaire en sortie d'étude.

La p-valeur de la variable « Taux de chômage régional » est égale à 0,9753 alors l'hypothèse  $H_0$  n'est pas rejetée. Le taux de chômage régional de la région où se situe le master n'a donc pas d'impact sur le salaire en sortie d'étude selon le modèle économétrique. Ce résultat ne correspond pas à ceux de Caro, Checcaglini et Guironnet (2022) qui montrent que les jeunes diplômés en France issus de région ayant un taux de chômage plus élevé que la moyenne ont des salaires plus faibles. Cependant, il correspond à ceux Nikulin (2015) qui montre qu'il n'y a pas de corrélation entre le salaire et le taux de chômage en Europe.

La variable « Taux d'emplois cadre » a une p-valeur inférieure à 0,0001 donc l'hypothèse  $H_0$  est rejetée au niveau 1%. Selon les résultats économétriques, pour un point de pourcentage supplémentaire d'emplois de cadre en sortie d'étude dans un master, le salaire en sortie d'étude de ce master augmente de 0,271%, toutes choses égales par ailleurs. Alors, le taux d'emplois de cadre en sortie d'étude d'un master a un impact positif sur le salaire en sortie d'étude. Cela correspond avec les résultats de la littérature : Charruau et Épaulard (2017) montrent en effet que la position de cadre dans une entreprise est un élément déterminant du niveau salarial d'un individu en France. Aeberhardt et Pouget (2006) sont du même avis : ils expliquent que le salaire est étroitement lié à la catégorie socio-professionnelle.

La variable « Académie » est divisée en 5 dummies ayant pour catégorie de référence la zone géographique Île-de-France. Les 4 autres dummies sont : « Nord-Est », « Nord-Ouest », « Sud-Est » et « Sud-Ouest ». Toutes les catégories ont une p-valeur inférieure à 0,0001 donc la variable domaine est significative au niveau 1% et l'hypothèse  $H_0$  est rejetée. Un diplômé d'un master au Nord-Est de la France a un salaire en sortie d'étude inférieur de 6,827% en moyenne, toutes choses égales par ailleurs, par rapport à un diplômé d'une université en Île-de-France. Un diplômé d'un master situé au Nord-Ouest de la France a un salaire en sortie d'étude inférieur de 9,903% en moyenne, toutes choses égales par ailleurs, par rapport à un diplômé d'une université en Île-de-France. Le salaire en sortie d'étude d'un étudiant diplômé d'un master au Sud-Est de la France est inférieur de 7,193% en moyenne, toutes choses égales par ailleurs, par rapport à celui

d'un étudiant en Île-de-France. Puis, le salaire en sortie d'étude des diplômés d'un master dans le Sud-Ouest de la France est inférieur de 9,976% en moyenne, toutes choses égales par ailleurs, à celui des diplômés d'un master en Île-de-France. La zone géographique Île-de-France est celle qui offre les plus hauts salaires en sortie d'étude à la fin d'un master. En effet, toutes les autres régions ont des salaires qui, en moyenne, sont inférieurs, toutes choses égales par ailleurs. L'académie a un impact sur le salaire en sortie d'étude. Ces résultats correspondent à ceux de la littérature : Charruau et Épaulard (2017) montrent que la région géographique joue un rôle dans la détermination des salaires. En effet, ils observent que les zones géographiques les plus dynamiques offrent des salaires plus élevés.

La variable « Domaine » est divisée en 5 catégories et la catégorie de référence est le domaine « Droit, économie et gestion ». Les 4 autres domaines sont : « Sciences, technologies et santé », « Sciences humaines et sociales », « Lettre, langue et art » et « Master enseignement ». La p-valeur est inférieure à 0,0001 pour toutes les catégories donc l'hypothèse  $H_0$  est rejetée. La variable « Domaine » est significative dans le modèle économétrique. Un étudiant diplômé d'un master en sciences humaines et sociales perçoit un salaire en sortie d'étude inférieur de 13,330% en moyenne, toutes choses égales par ailleurs, par rapport à un étudiant diplômé dans le domaine droit, économie et gestion. Le salaire en sortie d'étude dans le domaine sciences, technologies et santé est inférieur de 7,225% en moyenne, toutes choses égales par ailleurs, par rapport au salaire en sortie d'étude de la catégorie de référence. Le salaire en sortie d'étude du domaine lettre, langue et art est inférieur de 10,063% en moyenne, toutes choses égales par ailleurs, par rapport au salaire des diplômés du domaine droit, économie et gestion. Puis, le salaire en sortie d'étude du domaine master enseignement est inférieur de 22,611% en moyenne, toutes choses égales par ailleurs, par rapport au salaire des diplômés du domaine droit, économie et gestion. Le domaine droit, économie et gestion est le domaine qui offre les salaires en sortie d'étude les plus élevés, toutes choses égales par ailleurs. Le domaine a donc un impact sur le salaire en sortie d'étude. Ces résultats correspondent avec la littérature : Martinelli et Prost (2010) montrent que le domaine d'étude joue un rôle majeur dans la détermination des salaires en sortie d'étude en France. Aeberhardt et Pouget (2006) confirment ces résultats et les expliquent par des rendements du travail différent entre chaque domaine.

### **c) Conclusion**

Le modèle de régression linéaire Log-Level permet donc de mettre en évidence les effets, toutes choses égales par ailleurs, des différentes variables explicatives sur la variable dépendante : le salaire en sortie d'étude. Toutes les variables sauf le taux de chômage et le taux de diplômés boursiers sont significatives dans notre modèle. Le taux de femme, le taux d'insertion, le taux d'emplois de cadre, l'académie et le domaine ont un impact sur le salaire en sortie d'étude, toutes choses égales par ailleurs, et cela correspond avec la littérature.

### **3) Modèle de régression linéaire avec les effets croisés**

$$\begin{aligned} \text{Log}(\text{salnetmedian}) = & \beta_0 + \beta_1 \text{prfemme} + \beta_2 \text{tinsert} + \beta_3 \text{prdiplom} + \\ & \beta_4 \text{tchom} + \beta_5 \text{prcadre} + \beta_6 \text{ne} + \beta_7 \text{no} + \beta_8 \text{se} + \beta_9 \text{so} + \beta_{10} \text{shs} + \beta_{11} \text{sts} + \beta_{12} \text{lla} + \\ & \beta_{13} \text{prfemme\_shs} + \beta_{14} \text{prfemme\_sts} + \beta_{15} \text{prfemme\_lla} + \\ & \beta_{16} \text{prfemme\_me} + \beta_{17} \text{prcadre\_prfemme} + \varepsilon \end{aligned}$$

### a) Présentation du modèle

Les effets croisés ajoutés au modèle sont : l'effet entre la variable « Taux de femmes » et la variables « Domaine » et l'effet entre « Taux de femmes » et « Taux d'emplois cadre ». Dans la littérature, Battagliola (2001) montre que les femmes sont sous ou sur-représentées dans certains domaines. Sachant que le taux de femmes impacte négativement le salaire en sortie d'étude, comme vu précédemment, et que les salaires en sortie d'étude ne sont pas les mêmes selon les domaines, il est alors intéressant de mesurer cet effet croisé entre le taux de femmes dans un master et le domaine du master. De plus, Blau et Kahn (2017) précisent que la position dans la hiérarchie de l'entreprise est un des déterminants clefs de l'écart salariale entre les hommes et les femmes. Il est alors intéressant de croiser les variables « Taux de femmes » et « Taux d'emplois cadre ».

La significativité générale du modèle est testée. Tous les coefficients de régression sont testés simultanément : les hypothèses sont, H0 : « Tous les coefficients de régression sont nuls » et H1 : « Au moins un des coefficients de régression est non nul ». La statistique de test suit une loi de Fisher et la p-valeur est inférieure à 0,0001. L'hypothèse H0 est donc rejetée au niveau 1%. Ainsi au moins un des coefficients de régression est non nul et le modèle a au moins une variable explicative.

Le coefficient de détermination  $R^2$  est égale à 0,8058 ce qui signifie que 80,58% de la variance de la variable dépendante Y est expliquée par les variables explicatives du modèle. Une augmentation du  $R^2$  par rapport au modèle Log-Level est constaté, cependant comme des variables sont ajoutées à ce modèle, il est évident que celui-ci augmente. L'analyse du  $R^2$ -ajusté est donc plus intéressante. Le  $R^2$ -ajusté du modèle avec les effets croisés est égale à 0,8013 donc 80,13% de la variance de la variable dépendante Y est expliquée par les variables significatives du modèle. Le  $R^2$ -ajusté est supérieur à celui du modèle Log-Level, donc l'ajout des variables des effets croisés permet de mieux expliquer la variable dépendante qu'auparavant.

### b) Résultats et interprétations

Comme pour le modèle Log-Level, la significativité de chaque variable est testée à l'aide de la statistique de Student et la p-valeur. Les hypothèses pour chaque test sont H0 : « le coefficient de régression est nul » et H1 : « le coefficient de régression est non nul ».

Les variables qui correspondent à l'effet croisé entre la variable « Taux de femmes » et la variable « Domaine » sont représentés par 5 catégories et la catégorie de référence est le taux de femme croisé avec le domaine droit, économie et gestion. Les autres catégories sont : le taux de femmes croisé avec le domaine sciences humaines et sociales, soit la variables « *prfemme\_shs* ». Le

taux de femmes avec le domaine sciences, technologies et santé, soit la variables « *prfemme\_sts* ». Le taux de femmes avec le domaine lettre, langue et art, soit la variables « *prfemme\_lla* », et le taux de femmes avec le domaine master enseignement, soit la variable « *prfemme\_me* ». Seulement la variables « *prfemme\_sts* » a une p-valeur inférieure à 0,001. Les autres variables ont des p-valeurs supérieure à 0,10 donc l'hypothèse H0 n'est pas rejetée pour ces variables. Ainsi seulement la variables « *prfemme\_sts* » est significative. Cependant, comme une des 4 catégories est significative, il est alors intéressant d'étudier l'effet croisé entre le taux de femme et le domaine. Pour une augmentation d'un point de pourcentage du taux de femmes dans un master inclus dans le domaine sciences humaines et sociales, le salaire en sortie d'étude diminue en moyenne de 0,1878% pour ces masters, toutes choses égales par ailleurs. De même, pour une augmentation d'un point de pourcentage du taux de femmes dans un master inclus dans le domaine langue, lettre et art, le salaire en sortie d'étude diminue en moyenne de 0,312% dans ces masters, toutes choses égales par ailleurs. Puis, pour une augmentation d'un point de pourcentage du taux de femmes dans un master inclus dans le domaine master enseignement, le salaire en sortie d'étude diminue en moyenne de 0,0589% pour ces masters, toutes choses égales par ailleurs. Cependant, pour une augmentation d'un point de pourcentage du taux de femmes dans un master inclus dans le domaine sciences, technologies et santé, le salaire en sortie d'étude augmente en moyenne de 0,19% pour ces masters, toutes choses égales par ailleurs. Il y a effectivement une différence de salaire selon les domaines en fonction du pourcentage de femmes présentes dans le master. Cela correspond avec la littérature : Battagliola (2001) montre que les femmes ont tendance à être sous représentées dans certains domaines, ce qui peut exercer une influence sur le salaire en sortie d'étude. Il est observé à l'aide du modèle que le domaine sciences, technologies et santé est le domaine qui offre le meilleure salaire en sortie d'étude, pour un point de pourcentage supplémentaire du taux de femme.

La variable de l'effet croisé entre le taux de femmes et le taux d'emplois cadre en sortie d'étude correspond à la variable « *prfemme\_prcadre* ». Sa p-valeur est inférieure à 0,0001 donc l'hypothèse H0 est rejetée et la variable est significative au niveau 1%. Il existe donc un effet croisé entre le taux de femmes et le taux d'emplois cadre. Pour une augmentation d'un point de pourcentage du taux de femmes, à un taux d'emplois cadre fixé, le salaire en sortie d'étude diminue de  $(0,00137 + 0,00002991 * \text{« prcadre »}) * 100\%$ , toutes choses égales par ailleurs. Plus le taux d'emplois cadres est haut, plus la diminution du salaire est forte. Pour une augmentation d'un point de pourcentage du taux d'emplois cadre, le salaire en sortie d'étude augmente de  $(0,00486 - 0,00002991 * \text{« prfemme »}) * 100\%$ , toutes choses égales par ailleurs. Plus le taux de femmes est haut, moins l'augmentation du salaire est importante. Ainsi, une relation négative entre le taux de femmes et le taux d'emplois cadres est observée. Un master avec un haut taux d'emplois cadre et un haut taux de femmes aura des salaires en sortie d'étude moins élevés qu'un master avec un haut taux d'emplois cadre et un faible taux de femmes. L'étude économétrique montre que même dans les emplois cadre, les femmes ont des salaires moins élevés que les hommes. Ces résultats correspondent avec la littérature : Battagliola (2001) note que les femmes parviennent difficilement aux positions dirigeantes. Seulement un tiers des emplois cadres en France est occupé par des femmes.



### c) Conclusion

L'ajout des effets croisés dans le modèle de régression linéaire permet donc d'améliorer la qualité du modèle mais aussi de mettre en évidence les effets d'interaction entre les variables explicatives. Il est observé que l'impact sur le salaire en sortie d'étude était plus ou moins grand selon le domaine d'étude pour une augmentation du taux de femmes. Le modèle économétrique a aussi mis en évidence une relation négative entre le taux d'emplois cadre et le taux de femmes : plus un master a un taux d'emplois cadre élevé plus les salaires en sortie d'étude sont haut, mais plus le taux de femmes est élevé et plus les salaires sont faibles.

### 4) Modèle de régression linéaire sur 2 échantillons

$$\begin{aligned} \text{Log}(\text{salnetmedian}) = & \beta_0 + \beta_1 \text{prfemme} + \beta_2 \text{tinsert} + \beta_3 \text{prdiplom} + \\ & \beta_4 \text{tchom} + \beta_5 \text{prcadre} + \beta_6 \text{ne} + \beta_7 \text{no} + \beta_8 \text{se} + \beta_9 \text{so} + \beta_{10} \text{shs} + \beta_{11} \text{sts} + \beta_{12} \text{lla} + \\ & \beta_{13} \text{me} + \varepsilon \end{aligned}$$

#### a) Présentation du modèle

Le modèle de régression linéaire Log-Level est appliqué cette fois-ci sur 2 échantillons créés à partir de la base de données : un échantillon contenant toutes les observations ayant un salaire en sortie d'étude supérieur au salaire médian de la base de données, soit 1950€. Cet échantillon comprend 390 observations. Et un échantillon contenant les observations ayant un salaire en sortie d'étude inférieur ou égale au salaire médian, avec 420 observations. Cela peut permettre d'observer si le modèle fonctionne de la même manière sur la partie haute des salaires en sortie d'étude que sur la partie basse en comparant les coefficients de régression. L'échantillon contenant les observations ayant un salaire en sortie d'étude supérieur au salaire médian est appelé : « modèle higher » et l'échantillon contenant les observations ayant un salaire en sortie d'étude inférieur ou égale au salaire médian est appelé « modèle lower ».

La significativité générale des modèles est testée. Tous les coefficients de régression sont testés simultanément : les hypothèses sont,  $H_0$  : « Tous les coefficients de régression sont nuls » et  $H_1$  : « Au moins un des coefficients de régression est non nul ». La statistique de test du modèle higher suit une loi de Fisher et la p-valeur est inférieure à 0,0001. L'hypothèse  $H_0$  est donc rejetée au niveau 1%. Ainsi, au moins un des coefficients de régression est non nul et le modèle a au moins une variable explicative. Pour le modèle lower, la statistique de test suit une Fisher et la p-valeur est inférieure à 0,0001. L'hypothèse  $H_0$  est alors rejetée au niveau 1%. De ce fait, au moins un des coefficients de régression est non nul et le modèle a au moins une variable explicative.

Le coefficient de détermination ajusté :  $R^2$ -ajusté est égal à 0,5152 pour le modèle higher. Cela signifie que 51,52% de la variance de la variable dépendante Y est expliquée par les variables explicatives du modèle. Le  $R^2$ -ajusté du modèle lower est égal à 0,5944 ce qui est supérieur à celui du modèle higher. Ainsi, 59,44% de la variance de la variable dépendante Y est expliquée par les variables explicatives du modèle. Les variables explicatives expliquent donc d'une meilleure façon

la partie basse des salaires en sortie d'étude que la partie haute. Notons que le  $R^2$ -ajusté de chacun des modèles est plus faible que celui du modèle sur toutes les observations. Cela est dû au fait que l'échantillon étudié est plus petit.

## **b) Résultats et interprétations**

La significativité de chaque variable est testée à l'aide de la statistique de Student et de la p-valeur. Les hypothèses pour chaque test sont  $H_0$  : « le coefficient de régression est nul » et  $H_1$  : « le coefficient de régression est non nul ».

Tout d'abord, la variable taux de femmes a une p-valeur inférieure à 0,01 dans les deux modèles, alors l'hypothèse  $H_0$  est rejetée au niveau 1%. La variable est donc significative. Cependant, dans le modèle higher, une augmentation d'un point de pourcentage du taux de femmes dans un master entraîne une diminution du salaire en sortie d'étude de ce master de 0,0816%, toutes choses égales par ailleurs, tandis que dans le modèle lower cela entraîne une diminution de 0,115%, toutes choses égales par ailleurs. Une augmentation du taux de femmes a donc un impact plus important sur la partie basse des salaires en sortie d'étude que sur la partie haute. Cela peut être lié au fait que comme il est vu précédemment, le taux de femme est plus élevé dans les domaines ayant des salaires en sortie d'étude plus faible.

La p-valeur de la variable taux de chômage régional est de 0,0362, alors l'hypothèse  $H_0$  est rejetée au niveau 5%. La variable taux de chômage régional n'est donc pas significative dans le modèle Log-Level sur l'ensemble des observations, ni dans le modèle higher mais l'est dans le modèle lower. Le taux de chômage régional a un impact sur le salaire en sortie d'étude lorsque le salaire est inférieur ou égal au salaire médian. Dans ce cas, pour une augmentation d'un point de pourcentage du taux de chômage régional, le salaire en sortie d'étude des masters inclus dans la région en question augmente de 0,504%, toutes choses égales par ailleurs. Ces résultats ne correspondent pas avec la littérature. En effet, Caro, Checcaglini et Guironnet (2022) montrent que plus le taux de chômage est élevé dans une région, plus les salaires dans cette région sont faibles. Tandis que l'étude économétrique montre que plus le taux de chômage augmente, plus les salaires sont élevés dans le modèle lower.

La variable taux d'emplois cadre a un p-valeur inférieure à 0,01 dans les deux modèles, alors l'hypothèse  $H_0$  est rejetée au niveau 1%. Ainsi, la variable est significative dans ces modèles. Cependant, cette variable n'a pas le même impact sur le salaire en sortie d'étude en fonction de si le salaire est supérieur ou inférieur ou égal au salaire médian. Pour une augmentation du taux d'emplois cadre dans un master d'un point de pourcentage dans le modèle higher, le salaire en sortie d'étude de ce master augmente de 0,245%, toutes choses égales par ailleurs. Cependant, dans le modèle lower, pour une augmentation d'un point de pourcentage du taux de cadre dans un master, le salaire en sortie d'étude pour ce master augmente de 0,148%, toutes choses égales par ailleurs. Le taux d'emplois cadre d'un master a donc un impact plus important sur le salaire en sortie d'étude lorsque les salaires sont supérieurs au salaire médian que lorsque les salaires sont inférieurs ou égaux au salaire médian.

Pour la variable académie, le modèle économétrique ne montre pas de différence majeure entre la partie haute des salaires et la partie basse, toutes choses égales par ailleurs. Cependant, il y a des différences entre les coefficients de régression des domaines étudiés. Tout d'abord, il n'y a aucun master du domaine master enseignements dans le modèle higher. Aucun diplômé du domaine master enseignement à la sortie d'étude ne gagne plus que le salaire médian à la sortie d'étude qui est de 1950€. Dans le modèle lower, un diplômé du master enseignement a un salaire à la sortie d'étude inférieur de 11,252% en moyenne, toutes choses égales par ailleurs, à un étudiant diplômé du domaine droit, économie et gestion. Le domaine sciences humaines et sociales a une p-valeur inférieure à 0,01% dans les deux modèles. La variable est significative au niveau 1% autant pour la partie haute que pour la partie basse des salaires. De plus, la différence de variation du salaire en sortie d'étude lorsque le domaine est sciences humaines et sociales est faible. En effet, le salaire diminue en moyenne de 7,4%, toutes choses égales par ailleurs, dans le modèle higher et de 8,056% en moyenne, toutes choses égales par ailleurs, pour le modèle lower. Le domaine sciences humaines et sociales impacte donc de la même manière le salaire en sortie d'étude, que le salaire soit supérieur ou inférieur ou égale au salaire médian. Le domaine sciences, technologies et santé a une p-valeur inférieure à 0,01, donc la variable est significative au niveau 1% dans les deux modèles. Cependant elle n'impacte pas les salaires en sortie d'étude de la même manière. Effectivement, le salaire est inférieur de 5,805% en moyenne, toutes choses égales par ailleurs, dans le modèle higher et est inférieur de 2,409% en moyenne, toutes choses égales par ailleurs, dans le modèle lower, par rapport à un étudiant diplômé du domaine droit, économie et gestion. Donc les écarts de salaire entre le domaine sciences, technologies et santé et le domaine droit, économie et gestion sont plus importants dans la partie haute des salaires en sortie d'étude que dans la partie basse. La p-valeur du domaine langue, lettre et art est supérieure à 0,10 dans le domaine higher et est inférieure à 0,1 dans le domaine lower. Ainsi, la variable n'est pas significative dans le modèle higher mais l'est dans le modèle lower. Cela est lié au fait que dans le modèle higher, il y a seulement une observation avec le domaine langue, lettre et art. De ce fait, seulement un master du domaine langue, lettre et art a un salaire en sortie d'étude supérieur au salaire médian. Dans le modèle lower, un diplômé de ce domaine a un salaire inférieur de 11,252% en moyenne, toutes choses égales par ailleurs, à un diplômé du domaine droit, économie et gestion.

Ainsi, selon si le master a un salaire en sortie d'étude supérieur ou inférieur ou égale au salaire médian, le domaine ne va pas impacter le salaire en sortie d'étude de la même manière. Cependant, comme il est observé dans le modèle Log-Level sur l'échantillon complet, le domaine droit, économie gestion est toujours celui avec les plus hauts salaires en sortie d'étude.

### **c) Conclusion**

Ces deux modèles sur ces deux échantillons différents permettent de mettre en évidence que l'impact des variables n'est pas le même en fonction de si le salaire en sortie d'étude est inférieur ou égale, ou supérieur au salaire médian. En effet, le taux de femmes, le taux de chômage et le taux d'emplois cadre n'impactent pas le salaire en sortie d'étude de la même manière. Le taux de femmes joue un rôle plus important dans la partie basse des salaires. Inversement, le taux d'emplois cadre impacte plus le salaire dans la partie haute que dans la partie basse. Puis, le taux de chômage régional n'est pas significatif dans la partie haute, mais impacte le salaire en sortie d'étude

positivement dans la partie basse. De plus, le modèle économétrique permet d'observer que les écarts de salaires en sortie d'étude entre le domaine droit, économie et gestion et les autres domaines n'étaient pas les mêmes en fonction de si le salaire en sortie d'étude est inférieur ou égale, ou supérieur au salaire médian. En effet, l'écart de salaire entre le domaine sciences humaines et sociales et le domaine droit, économie gestion est plus fort dans la partie basse des salaires. Cependant, l'écart de salaires entre le domaine sciences, technologies et santé et le domaine droit, économie et gestion est plus fort dans la partie haute des salaires que dans la partie basse. Le domaine master enseignements n'a aucune observation avec un salaire supérieur au salaire médian et le domaine langue, lettre et arts en a seulement une. Les salaires en sortie d'étude de ces domaines sont donc inférieurs ou égaux au salaire médian.

## **Conclusion**

Le projet réalisé a pour objectif d'étudier les différences salariales à la sortie d'un master universitaire en France en 2017. Pour ce faire, la littérature économique a permis de sélectionner sept variables explicatives pouvant expliquer ces différences salariales.

L'analyse statistique montre que les variables explicatives les plus corrélées avec le salaire sont le taux de femmes et le taux d'emploi cadre. Comme dans la littérature, une forte relation empirique négative est observée entre le salaire à la sortie d'un master et le taux de femmes présent dans ce master, ainsi qu'une forte corrélation positive entre le taux d'emploi cadre et le salaire à la sortie d'étude. Aussi en accord avec la littérature, le salaire à la sortie d'un master est lié négativement au taux de chômage de la région de ce master. Le salaire à la sortie d'étude diffère notamment selon le domaine d'étude. En effet les salaires les plus élevés sont obtenus par les diplômés en droit, économie et gestion ainsi qu'en sciences, technologies et santé. Le salaire à la sortie d'un master varie en sens inverse du taux de boursier présent dans celui-ci. L'académie est déterminante quant aux salaires de ses diplômés. La zone géographique Île-de-France se distingue notamment en affichant un salaire net médian de ses diplômés plus élevé que ceux des quatre autres zones géographiques. Néanmoins, le lien obtenu à l'issue de l'étude statistique entre le salaire et le taux d'insertion diffère du résultat suggéré par la littérature. En effet, l'étude empirique montre une corrélation positive entre le taux d'insertion de l'établissement et le salaire perçu par ses diplômés, là où d'après la littérature, le taux d'insertion n'aurait pas d'effet sur le salaire.

En somme, l'analyse statistique a permis de démontrer qu'il existe bien des corrélations entre le salaire net médian mensuel des emplois à temps plein à la sortie d'un master universitaire en France en 2017 et les variables explicatives sélectionnées à l'aide de la littérature. Toutefois, ce type d'analyse présente la limite de ne pas pouvoir étudier le lien d'une variable explicative avec la variable expliquée, indépendamment des liens entre toutes les variables explicatives. Ne raisonnant pas toutes choses égales par ailleurs, le salaire peut être influencé par des facteurs externes. Pour ces raisons, une étude économétrique est réalisée.

Le modèle de régression linéaire multiple permet de mettre en évidence la significativité des variables citées précédemment. Plus précisément, le taux d'insertion, le taux d'emploi cadre, et l'académie apparaissent comme les variables les plus pertinentes du modèle. En effet, pour les trois sous modèles, ces variables restent significatives à 1%. Comme le prévoyait la littérature économique, les masters avec un taux d'emploi cadre plus élevé débouchent sur un salaire plus élevé que dans les autres masters. De la même façon, un master diplômant en Île-de-France donne accès à un revenu plus élevé qu'un master diplômant dans les autres zones géographiques. Pareillement, les masters inclus dans le domaine droit, économie et gestion présentent des salaires plus élevés que les autres. Concernant le taux de femmes, un master avec un taux plus élevé présente un salaire inférieur aux autres masters, ce qui concorde avec la littérature. Cependant, le taux d'insertion a un impact positif sur le salaire, ce qui s'oppose avec les résultats attendus d'après la littérature. Toujours à l'instar de la littérature, le taux de chômage et le taux de boursiers ne sont pas significatifs.

Le modèle croisé a permis d'améliorer la qualité du modèle, notamment en mettant en exergue des effets croisés entre les variables explicatives précédemment mentionnées par la littérature. À ce titre, l'étude économétrique souligne l'effet interactif entre le taux de femmes et le taux d'emploi cadre dans un master. Effectivement, plus un master présente un taux d'emploi cadre élevé plus le salaire associé est élevé. Néanmoins plus le taux de femmes est élevé dans ce dernier et plus le salaire est faible.

Pour finir, les modèles higher et lower ont permis d'approfondir cette étude en étudiant séparément les données associées à des salaires supérieurs et inférieurs au salaire médian. Notamment, le taux de chômage impacte négativement le salaire lorsque celui-ci se situe en-deçà de la médiane alors qu'il n'est pas significatif lorsque le salaire excède la médiane. De même, l'impact négatif du taux de femmes sur le salaire perçu à la sortie d'un master est plus important lorsque le salaire est inférieur à la médiane.

Cette étude présente plusieurs limites. Tout d'abord, la base de données repose sur un questionnaire. Ce faisant, le taux de réponse peut avoir son importance sur les résultats obtenus. Le taux de réponses représente le pourcentage de réponse donnée par les étudiants pour chacune des formations vis-à-vis de leur salaire. Le taux de réponse n'a pas de lien direct avec la variable d'intérêt mais peut justifier des disparités si les taux de réponses diffèrent considérablement entre les formations. D'autre part, l'étude comporte seulement sept variables explicatives. L'ajout d'autres variables indépendantes peut entraîner une augmentation de la variance expliquée par le modèle. Par exemple, Giret et Issehnane (2012) soulignent que le nombre de stages effectués durant la scolarité a un impact positif sur le premier salaire. En effet, ces derniers peuvent être interprétés comme un témoignage de la motivation de l'étudiant, ou encore comme une sorte de pré-recrutement pour les employeurs qui souhaitent employer leur stagiaire à l'issue de leur stage selon eux. De plus, cette étude ne traite que les master universitaires et ne prend pas en compte les masters issus d'autre cursus tel que ceux d'école d'ingénieur ou de commerce qui proposent un apprentissage différent selon Barth.

## **Bibliographie**

- Aaronson, Sullivan, 2000, « Unemployment and Wage Growth: Recent Cross-State Evidence ».
- Aeberhardt, Pouget, 2006, « Comment expliquer les disparités salariales ? » - Les salaires en France, édition 2006 – INSEE
- Barth « Master à l'université ou dans une grande école : comment choisir ? »
- Battagliola, 2001, « Le Travail des Femmes: Une Paradoxe Émancipation ».
- Blau, Kahn, 2017, « The Gender Wage Gap: Extent, Trends, and Explanations » - *Journal of Economic Literature* 55, no°3 (2017): 789–865.
- Bourdon, Giret, Goudard, 2011, « Peut-on classer les universités en fonction de leurs performances d'insertion ? » 2011, halshs-00564859.
- Brown, Roberts, Taylor, 2011 « The gender reservation wage gap: Evidence from British Panel data » - *Economics Letters* 113, no°1: 88-91.
- Caro, Checcaglini, Guironnet, 2022, « L'empreinte des territoires sur le salaire des jeunes débutants » - Chemin vers l'emploi et la vie adulte : l'inégalité des possibles, Céreq, pp.62-67, 2022, 978-2-11-151946-6, halshs-03936395.
- Charruau, Épaulard, Mai 2017, « D'où viennent les écarts de salaire entre les territoires ? », La note d'analyse, Mai 2017, n°55.
- Demeuse, Friant, Malaise, 2014, « Réussite à l'université et composition socioéconomique de l'établissement d'enseignement secondaire fréquenté » - *Une analyse de données longitudinale en Belgique francophone. XXIe journées d'étude sur les données longitudinales dans l'analyse du marché du travail. Réussite scolaire, réussite professionnelle, l'apport des données longitudinales*, Jun 2014, Dijon, France. pp.165-174.
- Gaini, Leduc, Vicard, 2013, « Peut-on parler de « générations sacrifiées » ? Entrer sur le marché du travail dans une période de mauvaise conjoncture économique. », In : *Économie et statistique*, n°462-463, 2013. pp. 5-23.
- Giret, Goudard Octobre 2010, « Mesure des effets établissements sur le salaire des diplômés du supérieur. », Net.Doc, n°72, 2010, 30 p.
- Giret, Issehnane 2012 « L'effet de la qualité des stages sur l'insertion professionnelle des diplômés de l'enseignement supérieur » Formation emploi, no°117 : 29-47.
- Goldin, Pekkala Kerr, Olivetti, Barth, 2017, «The Expanding Gender Earnings Gap: Evidence from the LEHD-2000 Census » - *American Economic Review*, 107 (5): 110-14.
- Laufer, 2005, « La construction du plafond de verre : le cas des femmes cadres à potentiel »
- Lhéritier, 1992, « Les déterminants du salaire », *Économie et statistiques*, n°257, septembre 1992, pp. 9-21
- Loprest, 1992, « Gender Differences in Wage Growth and Job Mobility » *The American Economic Review* 82, no°2: 526-532.
- Martinelli, Prost, 2010, « Le domaine d'étude est déterminant pour les débuts de carrière » - INSEE Première, n°1313.
- Nimer, 2016, « La Construction des Inégalités Sociales dans le Système Éducatif au Liban: Étude de cas d'un Groupe de Boursiers dans une Université d'Élite ».