

2024/2025



## Rapport final

---

**Existent-ils des inégalités de genre dans l'aide que les parents fournissent à leurs enfants?**

Lison Brochet

Félix Jenft

Diane Sainteny

Sous la supervision de :

M. Bertrand Garbinti

# Contents

<b>I Premiers chiffres sur l'aide des parents selon le sexe</b>	<b>4</b>
I.1 Comparaison de l'aide moyenne des parents selon le sexe . . . . .	4
I.2 Comparaison de la satisfaction moyenne de l'aide des parents selon le sexe	6
I.3 Quelles variables utiliser pour expliquer cette différence sexuée ? . . . . .	6
I.3.1 LASSO sur le niveau de satisfaction des jeunes vis-à-vis de l'aide fournie par les parents . . . . .	7
I.3.2 LASSO sur l'aide totale mensuel perçue par les jeunes et fournie par les parents . . . . .	8
I.3.3 Quelques statistiques descriptives sur les variables explicatives choisies	9
<b>II Les femmes sont-elles toutes choses égales par ailleurs davantage aidées par leurs parents ?</b>	<b>12</b>
II.1 Influence des différentes variables de contrôle et conclusion sur l'impact du sexe sur l'aide des parents . . . . .	14
II.2 Limites du modèle Tobit I pour mesurer les marges intensives et extensives	17
<b>III Un écart sexué qui dépend du type d'aide</b>	<b>17</b>
III.1 Types d'aides parentales : niveau et fréquence . . . . .	17
III.2 Inégalités d'accès à l'aide parentale selon le sexe et le type d'aide . . . . .	18
III.2.1 Inégalité genrée d'accès à l'aide parentale pour le logement . . . . .	18
III.2.2 Inégalité genrée d'accès à l'aide parentale pour les loisirs . . . . .	20
III.3 Ecarts de montant des aides reçues entre les hommes et les femmes . . . . .	21
III.3.1 Ecarts de montant de l'aide reçue pour le logement entre les hommes et les femmes . . . . .	22
III.3.2 Ecarts de montant de l'aide reçue pour les loisirs entre les hommes et les femmes . . . . .	22
<b>IV Annexe</b>	<b>26</b>
IV.1 Opérations sur la base de données . . . . .	26
IV.2 Questionnements méthodologiques sur le LASSO . . . . .	27
IV.3 Construction de tranches de revenus . . . . .	27
IV.4 L'enjeu de prendre une catégorie de référence dans nos régressions . . . . .	29
IV.5 Choix de l'hétéroscédasticité et de la non utilisation de clusters . . . . .	29
IV.6 Motivation du modèle logit et effets marginaux/pourquoi logit plutôt que probit . . . . .	29
IV.7 Modèle tobit et effets marginaux . . . . .	29
IV.8 La décomposition de Gelbach . . . . .	31
IV.9 Le choix des types d'aide à étudier . . . . .	32
IV.10 Suite des régressions logistiques par type d'aide . . . . .	35
IV.11 Informations complémentaires sur l'étude menée sur les écarts dans les montants reçus par les jeunes . . . . .	38

## Introduction

L'INSEE calcule pour 2024 un écart de revenus entre les hommes et les femmes de 13,9% en équivalent temps plein. Beaucoup d'études s'intéressent à cet écart dans le milieu du travail. Toutefois, beaucoup moins s'interrogent sur le moment dans la vie des jeunes adultes à partir duquel un écart en matière d'apports financiers pourrait se faire sentir. Au tout début de l'âge adulte, alors que les ressources financières et le niveau de vie dépendent plus ou moins des aides parentales, un écart est-il déjà observé ? Et s'il y a en effet des inégalités genrées dans les transferts intrafamiliaux provenant spécifiquement des parents, favorisent-elles aussi les jeunes garçons au détriment des jeunes filles ?

## Revue de littérature

Ces enjeux ont déjà fait l'objet d'études antérieures, surtout aux États-Unis. Le fait que ces études reposent majoritairement sur des données provenant d'un autre pays – de surcroît non européen – doit être pris en compte. Il n'est en effet pas possible d'en transposer directement les conclusions au contexte français, dans la mesure où des différences culturelles, socio-économiques et normatives (tant au sens juridique que social) peuvent fortement influencer les modalités de l'aide financière apportée par les parents à leurs enfants.

Parmi les travaux les plus proches de notre problématique figure l'étude de McGarry et al., intitulée “*Transfer Behavior in the Health and Retirement Study: Measurement and the Redistribution of Resources Within the Family.*” (1995). Cette étude se base sur une enquête du Health and Retirement Study (HRS), qui ne concerne donc que les États-Unis. Les auteurs de cet article cherchent à savoir quels facteurs influencent l'aide que fournissent les parents à leurs enfants. Ils mettent en évidence que les parents ajustent leur aide en fonction des besoins de leurs enfants, privilégiant ceux qu'ils estiment les plus en difficulté. Cependant, à niveau de besoin égal, ils observent que les filles reçoivent en moyenne davantage d'aide – ou plus fréquemment – que les garçons, ce qu'ils interprètent comme le reflet d'une perception selon laquelle les filles auraient davantage besoin de soutien. Enfin, dans les familles les plus aisées, il semble que les garçons soient davantage soutenus pour financer leurs études ou l'achat d'un logement, tandis que les filles reçoivent plus fréquemment des dons monétaires. Ainsi, même à montant égal reçu, il existe une différenciation sexuée sur le type d'aides reçues.

Par ailleurs, certaines études se sont spécifiquement intéressées au soutien financier apporté par les parents pour les études supérieures de leurs enfants, comme Henretta et al. (2012) dans “*Family Structure and the Reproduction of Inequality: Parents' Contributions to Children's College Costs.*”. Cet article, centré sur les États-Unis, vise à analyser dans quelle mesure les parents participent au financement des études universitaires de leurs enfants, et à déterminer si cette contribution dépend principalement des caractéristiques individuelles des enfants, telles que le sexe, ou de celles de la structure familiale (par exemple, les familles recomposées). L'article montre que les parents investissent légèrement plus dans les jeunes hommes s'ils pensent que ceux-ci ont des meilleures perspectives salariales. Cependant, lorsque les jeunes femmes réussissent bien leurs études, comme en témoignent un GPA élevé ou l'obtention de bourses, elles bénéficient d'un soutien financier égal voire supérieur à celui de leurs homologues masculins. En résumé, les par-

ents investissent davantage dans l'enfant qu'ils considèrent comme le meilleur placement, c'est-à-dire celui qui tirera le plus profit de cet investissement. Toutefois, leur perception du bon placement diffère selon le sexe : pour les filles, cela correspond à de bonnes performances scolaires, tandis que pour les garçons, c'est plutôt la perspective d'un salaire élevé.

D'autres études abordent également cette problématique, bien que ce ne soit pas leur thème principal. On pense notamment à "*Only as Happy as the Least Happy Child: Multiple Grown Children's Problems and Successes and Middle-Aged Parents' Well-Being*" de Fingerman et al (2021). Les auteurs cherchent à établir un lien entre les accomplissements ainsi que les trajectoires des jeunes adultes et le bien-être des parents. Les parents semblent être en moyenne plus proches de leurs filles et plus au courant des difficultés auxquelles elles se trouvent confrontées. Cette perception se traduit par un soutien renforcé, tant matériel que non matériel. Les auteurs montrent que les parents ont tendance à soutenir prioritairement l'enfant qui semble le plus en difficulté, notamment lorsqu'il traverse une période compliquée, qu'elle soit d'ordre financier ou personnel. Cependant, en raison de biais genrés dans la perception, cet enfant est souvent la fille de la fratrie. Ainsi, cet article rejoint les conclusions du premier article mentionné, bien que leurs points de départ soient différents.

## Données

Pour répondre à ces interrogations, nous allons utiliser l'***Enquête Nationale sur les Ressources des Jeunes (ENRJ) - 2014***. Cette enquête est réalisée via questionnaire sur un échantillon représentatif de la population française. Le but de cette enquête est de recenser l'ensemble des ressources financières des jeunes adultes de 18 à 24 ans et leur provenance, tout en prenant en compte la situation de ces jeunes comme leur situation scolaire ou professionnelle, ainsi que l'état et les caractéristiques du logement dans lequel ils vivent. Trois questionnaires ont été élaborés dans le cadre de cette enquête. Le premier, intitulé QMENAGE, était à destination des unités d'habitation, à partir du moment où un des habitants était dans la tranche d'âge enquêtée. Les questionnaires QJEUNE et QPARENT ont été, eux, posés respectivement à des jeunes âgés d'entre 18 et 25 ans et à leurs parents.

Dans le cadre de notre projet, nous exploitons principalement les données des bases QJEUNE et QPARENT, bases dont la complémentarité va nous être très utile puisqu'elle permet notamment de comparer les perceptions des jeunes et des parents sur l'aide respectivement qu'ils reçoivent et qu'ils fournissent. L'existence de ces deux bases permet également de comparer les aides reçues selon le genre de l'enfant tout en contrôlant certaines caractéristiques des parents. Tous les milieux n'adoptent en effet certainement pas les mêmes pratiques en ce qui concerne l'existence et les modalités du soutien matériel et psychologique à leurs enfants – d'autant plus qu'il ne s'agit pas de vouloir aider tant son enfant mais aussi de le pouvoir, puisqu'aider coûte du temps et de l'argent—. Ainsi, comparer les aides reçues selon le genre sans contrôler un certain nombre de caractéristiques propres au jeune mais aussi à sa famille ne nous paraît pas pertinent.

## Question sur le sujet

Nous avons focalisé notre analyse sur les transferts monétaires directs et quantifiables des parents, type de soutien auquel on pense généralement exclusivement lorsqu'on évoque les aides parentales .

Si l'on s'en tient à cette dimension monétaire quantifiable, il est important de distinguer deux leviers par lesquels peut s'opérer une différenciation genrée :

- **à la marge extensive** : Est-ce que les garçons sont plus nombreux à recevoir une aide de leurs parents que les filles (ou inversement) ?
- **à la marge intensive** : Est-ce que les garçons qui reçoivent une aide de leurs parents reçoivent plus d'aides que les filles qui reçoivent une aide de leurs parents (ou inversement) ?

Au cours de notre travail, nous avons seulement étudié ce type d'aide. Il faut toutefois garder à l'esprit que le terme d'"aides" ne se limite pas à la notion de soutien financier. Les parents peuvent aussi aider leurs enfants en leur donnant accès à leur réseau, en les hébergeant, en les soutenant mentalement. S'il est beaucoup plus facile de mesurer et de quantifier l'aide monétaire, d'autres types d'aides sont aussi évalués dans le questionnaire et nous pensons qu'il est tout autant nécessaire de les étudier afin de mieux comprendre les causes potentielles d'inégalités genrées en matière d'aides parentales, bien que nous n'ayons pas pu le faire ici.

De plus, les aides monétaires sont elles-mêmes de natures très diverses. Ce n'est pas la même chose de donner de l'argent à son enfant pour l'aider à payer son logement que de lui donner de l'argent pour ses loisirs, ou pour l'entretien de sa voiture, voire pour en acheter une. Nous avons ainsi déjà regroupé toutes les variables d'aides en sept catégories pour pouvoir étudier de manière différenciée les différents types de soutien: les aides pour le logement, les transports, les loisirs, le numérique ( téléphone et connexion internet), la santé, la scolarité, ainsi que les aides à l'emprunt.

Pour élucider ces deux questions centrales de la marge intensive et de la marge extensive, d'autres facteurs mesurés dans l'enquête entreront également en jeu, comme la situation professionnelle, matrimoniale, ou familiale du jeune ou encore le revenu des parents.

## I Premiers chiffres sur l'aide des parents selon le sexe

### I.1 Comparaison de l'aide moyenne des parents selon le sexe

Nous avons d'abord réalisé une régression linéaire de la variable `aide_tot`, qui représente le montant mensuel total des aides financières parentales, sur une variable indicatrice du sexe (valant 1 pour une jeune femme), afin d'explorer les premières tendances du jeu de données.

Nous observons que les filles reçoivent en moyenne 40.3 euros d'aides en plus de la part de leurs parents que les jeunes garçons lorsque l'on prend en compte la marge extensive (modèle tobit), résultat significatif statistiquement (p-value quasi-nulle). Quand on ne prend pas en compte la censure du jeu de données en 0, cet écart est de 27.327€ en moyenne. L'estimation plus élevée du coefficient dans le modèle Tobit s'explique par le fait que ce modèle reconstitue les niveaux d'aide parentale comme s'ils n'étaient pas censurés, incluant ainsi des valeurs théoriquement possibles en dessous de zéro. Plus la

valeur associée à un jeune adulte est négative, plus la probabilité qu'il reçoive une aide financière de ses parents diminue. Le coefficient Tobit reflète à la fois les inégalités brutes en matière de revenus et la distance latente des jeunes non aidés par leurs parents vis-à-vis de la zone d'éligibilité à l'aide. Il est nécessaire de prendre en compte cette marge extensive via l'effet marginal calculé avec le modèle logit en plus de la marge intensive car son importance est loin d'être négligeable: une jeune femme a 27% de chance de plus de recevoir un montant non nul d'aides de ses parents qu'un jeune homme.

Table 1: L'écart moyen d'aide par les parents selon le sexe

	Tobit	OLS	Marge intensive	Effet marginal Logit
intercept 1	93*** (0.2)	166.413*** (5.045)	240.345*** (6.265)	0.1883*** (0.088)
sexe	40.3*** (0.3)	27.327*** (7.176)	23.817*** (8.784)	0.2782** (0.058)
Observations	5197	5197	4113	5197
Multiple $R^2$		0.00268	0.001785	
Ajusted $R^2$		0.002495	0.001542	

Notes: Les erreurs standards sont indiquées entre parenthèses.

\*\*  $p < 0.01$ , \*  $p < 0.05$ , •  $p < 0.1$ , •  $p < 0.5$ .

Cependant, ne regarder que la différence moyenne entre les jeunes femmes et les jeunes hommes ne permet pas de comprendre d'où vient la différence. La répartition du montant total d'aides parentales financières touché par les jeunes femmes est-elle similaire à celle des jeunes hommes, en étant décalée de 27€ ? Ou les filles sont-elles seulement sous-représentées ou sur-représentées dans certaines catégories de montant d'aides ce qui explique cet écart moyen final ? C'est ce que nous avons essayé de voir avec le tableau suivant.

aide	Femme	Homme
0€	45.9%	54.1%
0-5€	55.2%	44.8%
5-30€	47.8%	52.2%
30-62€	50.3%	49.7%
62-115€	47.1%	52.9%
115-180€	50.8%	49.2%
180-310€	52.2%	47.8%
310-570€	51.1%	48.9%
Plus de 570€	56.3%	43.7%

On voit que ce sont surtout dans les catégories les plus extrêmes que l'écart entre jeunes garçons et filles se creuse, c'est-à-dire chez les jeunes qui ne touchent aucune aide – un peu moins de 30% des jeunes – et ceux qui touchent plus de 570€ d'aides de leurs parents – 10% des jeunes –. On peut penser que cela découle en partie du fait que les femmes sont moins nombreuses que les hommes à arrêter leurs études très tôt et donc à être totalement indépendantes avant 24 ans.

De même, les jeunes femmes sont plus nombreuses que les hommes à atteindre un niveau d'étude supérieure au master, ce qui pourrait aussi expliquer qu'elles représentent 56,3% des jeunes recevant plus de 570€ de leurs parents.

Par conséquent, il est absolument nécessaire afin de comprendre cet écart de moyenne de 27€ par mois d'étudier le niveau de l'aide des parents en utilisant des variables de contrôle, comme le niveau d'éducation, mais aussi des indicateurs du milieu social des parents.

	satisfaction	total	total_individus	Femme	Homme
1	Autre	712524.2	765	26.79751	0.00000
2	Pas satisfait	1508072.2	1474	56.71747	0.00000
3	Satisfait	438323.6	472	16.48502	0.00000
4	Autre	689786.5	840	0.00000	25.30660
5	Pas satisfait	1507308.3	1614	0.00000	55.29950
6	Satisfait	528623.2	611	0.00000	19.39391

Étant donné l'écart moyen observé entre les jeunes hommes et les jeunes femmes en termes de montant d'aides parentales, il est légitime de se demander si cet écart se reflète également dans le niveau de satisfaction. Les hommes ont-ils conscience d'être moins aidés, et en sont-ils mécontents ? Dans ce tableau, ce que nous avons appelé satisfaction correspond en réalité à la perception du jeune quant à la possibilité que ses parents devraient l'aider davantage financièrement : oui (satisfait de ce qu'il a déjà) ou non (pas satisfait). Nous remarquons qu'un léger plus haut pourcentage de femme que d'hommes pensent que leurs parents pourraient les aider plus.

## I.2 Comparaison de la satisfaction moyenne de l'aide des parents selon le sexe

Un autre résultat intéressant, lisible dans le tableau ci-dessus, est que la plupart des hommes comme des femmes (plus de 50% dans les deux cas) pensent que leurs parents pourraient les aider plus financièrement.

## I.3 Quelles variables utiliser pour expliquer cette différence sexuée ?

Nous allons donc chercher quelles variables peuvent expliquer la différence de satisfaction des jeunes à l'égard des aides qu'ils reçoivent. Pour se faire nous avons recours à un LASSO. Par ailleurs nous avons fait de même avec le montant d'aides perçues en faveur des jeunes femmes et potentiellement compresser cette différence. Ce LASSO va éventuellement nous permettre de sélectionner les principales variables explicatives d'AIDETOT (aide mensuel donné par les parents aux jeunes). Nous présentons en annexe comment nous avons procédé pour la sélection des variables explicatives incluses dans chacun des LASSO. Par ailleurs, nous précisons que nous n'avons pas changé le nom des variables dans les tables LASSO car celles-ci sont parfois compliquées. C'est pourquoi les significations de toutes variables importantes est expliqué dans l'interprétation du LASSO, les variables n'étant pas détaillé sont d'une utilité réduite (par exemple CUMUL<sub>B</sub> est la manière dont les produits bancaires se sont accumulés. Si c'est en empruntant, la réponse est 1=oui, sinon 2. Un autre exemple est DISCVIEQUO<sub>M</sub> qui correspond au fait ou non de pouvoir parler de sa vie quotidienne à sa mère).

Table 2: Table des coefficients sélectionnés par Lasso avec  $R^2$  pour expliquer NIVAIDPM

	Coefficient
SOMREG <sub>C</sub>	-0.943381
AIDETOT	-0.064957
VET	-0.054523
CTRAIDFIM	0.045255
STIGTOT	-0.042255
FINMOIS	0.038288
JUGREV	-0.038224
TRANSPAR	0.035732
RPLUSRELA <sub>MB</sub>	-0.034861
MPRESAFTR	-0.031867
DISCVIEQUO <sub>M</sub>	0.028606
PREMSALPART	0.027068
TYPACT3	0.024919
ATRSP	-0.024881
CUMUL <sub>B</sub>	0.021557
FA	0.019592
NIVRESSPM	-0.018938
SORTI	-0.018814
JUGAID	0.018621
VOY	-0.015525
	$R^2 = 0.760$

Notes: NIVAIDPM=niveau de satisfaction du jeune à l'égard de l'aide financière fournie par ses parents

### I.3.1 LASSO sur le niveau de satisfaction des jeunes vis-à-vis de l'aide fournie par les parents

Dans ce LASSO, nous remarquons d'une part un  $R^2$  élevé ce qui signifierait que l'ensemble des variables incluse explique environ à 76% le niveau de satisfaction des jeunes. Il est aussi important de préciser que plus le niveau de satisfaction du jeune est bas plus la variable NIVAIDPM que l'on souhaite expliquer est élevé (1=tout à fait suffisante, 2=suffisante, etc... jusqu'à 8=refus et on exclu ceux qui ne savent pas) et que l'interpretation des signes de chaque coefficient est directe lorsqu'on prend en considération la définition de chaque variable.

Dans un premier temps, nous notons que la variable avec le coefficient, en valeur absolu, le plus important est SOMREG<sub>C</sub> (Est-ce que le jeune perçoit une aide financière régulière de la part de ses parents? Réponse : 1=ooui, 2=non) et non AIDETOT qui correspond à la somme perçu. La satisfaction du jeune à l'égard de l'aide perçue est donc plus déterminé par la régularité de l'aide que son montant.

On remarque que le sentiment de privation influe aussi sur ce niveau de satisfaction à travers des variables comme VET (sentiment de devoir souvent se priver de vêtement), VOY (sentiment de devoir souvent se priver de voyages ou de vacances) ou encore FIN-MOIS (le jeune adulte peut mettre de l'argent de côté à la fin du mois).

Par ailleurs, la perception que le jeune a du niveau de vie du ménage, donc de ses parents (FA) ce qui est lié avec le fait de devoir aider ou non ses parents financièrement (TRANSPAR) ont un impact compréhensible sur la satisfaction du jeune.



La présence d'autres variables est moins facilement explicable comme CTRAIDEFIM (part de l'aide prise en charge par la mère) ou RPLUSRELA<sub>MB</sub> (conflit avec la mère). En effet pourquoi la mère aurait un rôle plus important dans la satisfaction du jeune que le père? Peut-être est-ce parce que les jeunes femmes sont plus affectées par un conflit avec leur mère que pourrait l'être les jeunes hommes par rapport à l'un des deux parents? Cela permettrait d'expliquer en partie la disparité noté plus haut mais cette hypothèse ne peut rester qu'au stade d'hypothèse.

De plus, STIGTOT (la satisfaction du jeune sur de la vie qu'il mène) intervient aussi dans ce LASSO. Le jeune est-il non satisfait de sa vie à cause de l'aide qu'il perçoit? Ou est-ce parce qu'il est non satisfait de sa vie qu'il aurait potentiellement une mentalité négative et donc serait difficilement satisfait par le niveau d'aides perçues? On note un potentiel problème d'endogénéité avec cette variable. Cependant lorsque l'on retire cette variable et qu'on relance le LASSO, le  $R^2$  s'améliore très légèrement (0.761) et les autres coefficients varient très peu (à la troisième décimale uniquement).

Enfin la principale observation que l'on peut faire à l'aide de ce LASSO est qu'à aucun moment, la variable SEXEJ n'apparaît. Ainsi le niveau de satisfaction de l'aide n'est pas lié au fait d'être une fille ou un garçon, toutes choses égales par ailleurs. Cependant, l'on pourrait aussi dire que beaucoup de ces variables explicatives sont des variables de perception ou de sentiment du jeune et non des variables objectives mesurables. Ainsi l'on pourrait faire l'hypothèse, qui serait dure à prouver ou infirmer, que la manière de percevoir sa condition financière ou celle de ses parents, à quel point l'aide financière palie au manquement matériel, la sensation de se priver et ainsi de suite seraient différents entre les jeunes hommes et les jeunes femmes. Un corollaire de cette hypothèse pourrait être que les besoins ne seraient pas les mêmes entre les jeunes femmes et les jeunes hommes, que les jeunes femmes auraient des exigences plus élevées.

### **I.3.2 LASSO sur l'aide totale mensuel perçue par les jeunes et fournie par les parents**

L'interprétation de ce LASSO est plutôt directe. Nous remarquons que certaines variables que nous n'avions pas envisagées jusqu'à maintenant sont présentes. Par exemple, la variable COHAB (cohabitation avec les parents) semble jouer un rôle significatif dans le montant d'aide fourni par les parents. Cela est logique car si le jeune cohabite avec ses parents, ceux-ci n'ont pas besoin de lui fournir de l'aide pour le logement, peuvent éventuellement fournir une aide moindre pour la nourriture, etc...

Le fait d'être en décrochage scolaire ou le nombre d'heure travaillé semblent aussi jouer aussi un rôle prépondérant. Or les proportions d'hommes en décrochage ou qui travaillent sont plus élevées que pour les femmes et ne permet donc pas d'expliquer l'écart d'aide en faveur des femmes puisque le décrochage ou le nombre d'heures travaillées sont censées augmenter l'aide reçu.

Il en est de même pour la variable couple. Lorsqu'on régresse AIDETOT sur couple, on obtient un faible  $R^2$  (0.018) et être en couple diminue l'aide reçue uniquement si le jeune est en couple et habite avec son compagnon (COUPLE=1). Or la proportion d'homme dans cette situation est supérieure à la proportion de femme dans cette situation. On ne peut donc pas utiliser cette variable comme explication pour l'écart d'aides perçus entre hommes et femmes.

Table 3: Table des coefficients sélectionnés par Lasso avec  $R^2$  pour expliquer AIDETOT

	Coefficient
COHAB	74.102367
REVENU <sub>R</sub>	64.970268
DECRO	-63.270409
HEURTRAV <sub>A</sub>	-49.949850
COUPLE	29.543339
ENFANT	-19.914289
SEXEJ	10.728192
NIVET	-8.514966
SATRELP	7.534565
SITPRINC	5.582380
AGEJ	-4.845207
FORCHOIXANTD	-3.931220
STATUT	3.880672
RELATIONM	-3.706356
NPIECE1	3.673260
SPORT	3.136148
LOYER1 <sub>R</sub>	2.696317
NAT <sub>A</sub>	2.594029
SATRELM	2.446518
AGESEPARATIONP	2.142399
SURFACE1	1.980323
RELATIONP	-1.667813
AGEDECESP	-0.405503
	$R^2 = 0.355$

Notes: AIDETOT=aide financière totale fournie par les parents au jeune mensuellement

Plus généralement, ce LASSO n'est pas très concluant. Certaines variables ont un fort coefficient alors qu'une fois intégré dans la regression, leur coefficient est bien plus faible. Par ailleurs, l'on a accès à aucune des significativités de chaque coefficient ce qui rend le résultat peu rigoureux. Ainsi nous faisons le choix de faire une étude approfondie sur des variables présentes dans le LASSO et qui font intuitivement sens dans l'étude des interactions entre aide parentale au jeune. Cela nous permet d'être plus précis sur les variables sélectionnées et leur relation avec SEXEJ, même si pour cela l'on doit sacrifier l'exhaustivité. Plus spécifiquement, ces variables sont REVENU<sub>R</sub>, SITPRINC, auxquels nous ajoutons COHAB et AGEJ.

### I.3.3 Quelques statistiques descriptives sur les variables explicatives choisies

#### REVENU<sub>R</sub>

Pour analyser cette variable, il est intéressant d'avoir recours à des catégories de revenu (voir annexe pour leur construction) et d'avoir recours à des graphes pour se donner une idée de la répartition de nos données.

Avant de commenter ce graphique, il est important de préciser que les aires sous chaque courbe additionnées sont égales à 1. Ainsi l'aire sous chaque courbe représente la proportion de jeune dans chaque catégorie de revenu. Ce que l'on peut surtout retenir de ce graphique, c'est surtout que le pic de chaque densité se déplace sur la droite à partir de la quatrième catégorie et que le poids des valeurs à droite du pic deviennent plus importantes. Ceci peut sûrement s'expliquer par l'écart-type grandissant, comme vu ci-dessus.

Ce qu'on peut principalement soutirer de ce tableau, c'est avant tout que l'écart d'aides

Moyenne	Médiane	Écart-type
36.217826	15.000000	49342.679798
94.979107	44.000000	29774.050946
139.418121	89.000000	51229.689361
212.273477	150.000000	79572.975299
365.192440	333.000000	130143.671620
465.318563	416.000000	169731.783797

Figure 1: Table des aides par catégorie de revenu

On a affaire ici à des résultats très logiques. Plus les revenus des parents sont élevés, plus ils aident leur enfants. Ce qui est notable est l'écart-type grandissant qui peut s'expliquer par la répartition des populations dans chaque classe de revenu : plus il y a de parents dans une classe, plus l'écart-type est faible. Cet écart-type grandissant peut aussi s'expliquer par le fait que plus les parents gagnent des revenus importants, plus l'intervalle d'aide qu'ils peuvent fournir s'aggrandit, et donc plus il y a de possibilités de montant d'aides.

Catégorie	Moyenne		Médiane		Écart-type	
	Femmes	Hommes	Femmes	Hommes	Femmes	Hommes
Moins de 500	30.91	43.38	20.00	0.00	8282.26	96143.99
500 à 2000	106.82	82.23	48.00	40.00	32268.76	27454.07
2000 à 3000	148.15	131.04	90.00	86.00	55664.74	47280.25
3000 à 5000	235.37	190.90	173.00	132.50	86832.96	72441.29
5000 à 8000	396.32	332.53	365.00	303.00	135906.82	123957.74
Plus de 8000	597.76	380.74	561.67	370.00	242266.51	102084.12

Table 4: Statistiques par catégorie et sexe

genrées est en faveur des hommes pour la classe de revenus la plus basse et que cet écart s'inverse à partir de la deuxième classe de revenus et s'aggrandit au fur et à mesure qu'on monte dans les classes de revenus. Cette tendance est compliquée à expliquer sans prendre en compte d'autres variables (Plus les parents gagnent, plus leur filles font des études ou peut-être restent-elles plus longtemps en cohabitation avec leurs parents que leurs pairs masculins?) ou que par des suppositions qui sortent du cadre de cette étude. D'un autre côté, l'évolution des écarts-types peut s'expliquer par les effectifs (par exemple, pour la dernière catégorie, il y a deux fois moins d'effectifs pour les femmes que pour les hommes) et l'explication donnée ci-dessus (plus les parents gagnent, plus le choix de montant est élevé).

Ces graphiques permettent surtout d'illustrer qu'à la marge extensive (le fait de recevoir une aide ou non), les jeunes hommes semblent moins aidés que les jeunes femmes sauf pour la catégorie de revenu la plus basse.

## SITPRINC

De cette table, on retient que les jeunes en emploi sont les moins aidés, ce qui est logique car ceux-ci sont déjà indépendants en quelque sorte, alors que les jeunes en études ou en formation sont les plus aidés, ce qui n'est pas surprenant non plus. Les jeunes en apprentissage ou en formation constituent une catégorie entre-deux. Les résultats quand on regarde la différence entre les sexes sont plus surprenantes. Les jeunes hommes sont plus aidés

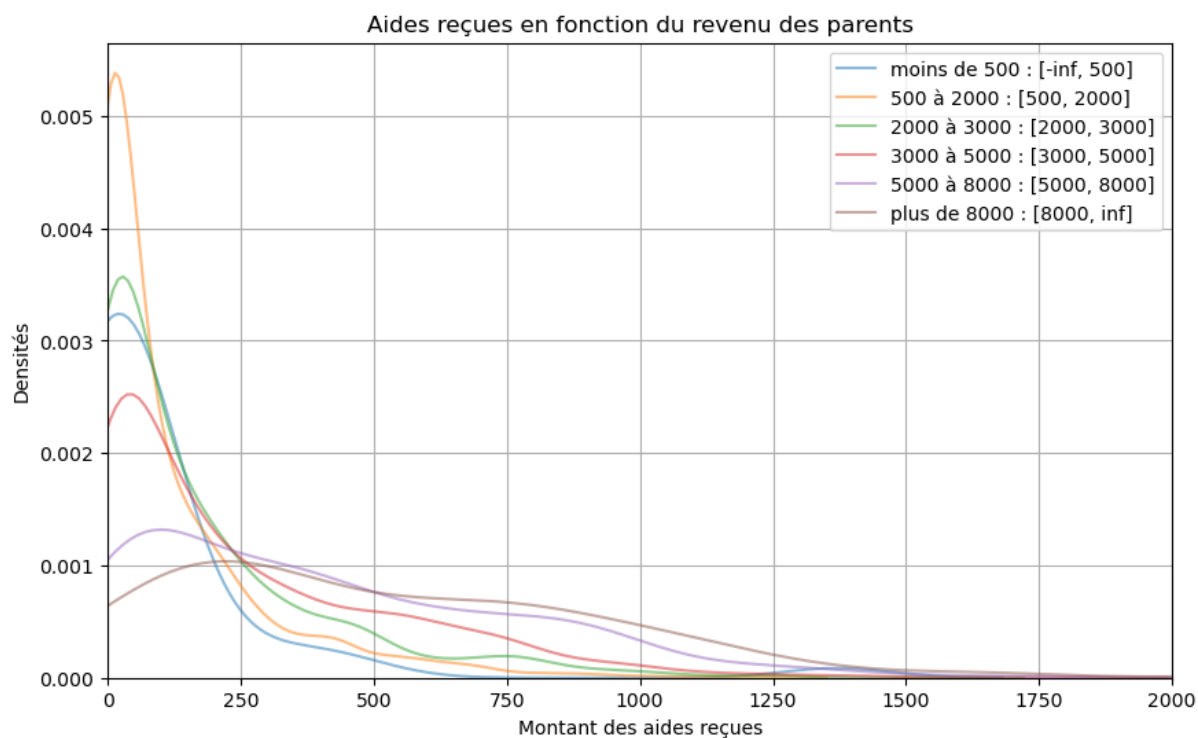


Figure 2: Distribution des aides en fonction du revenu

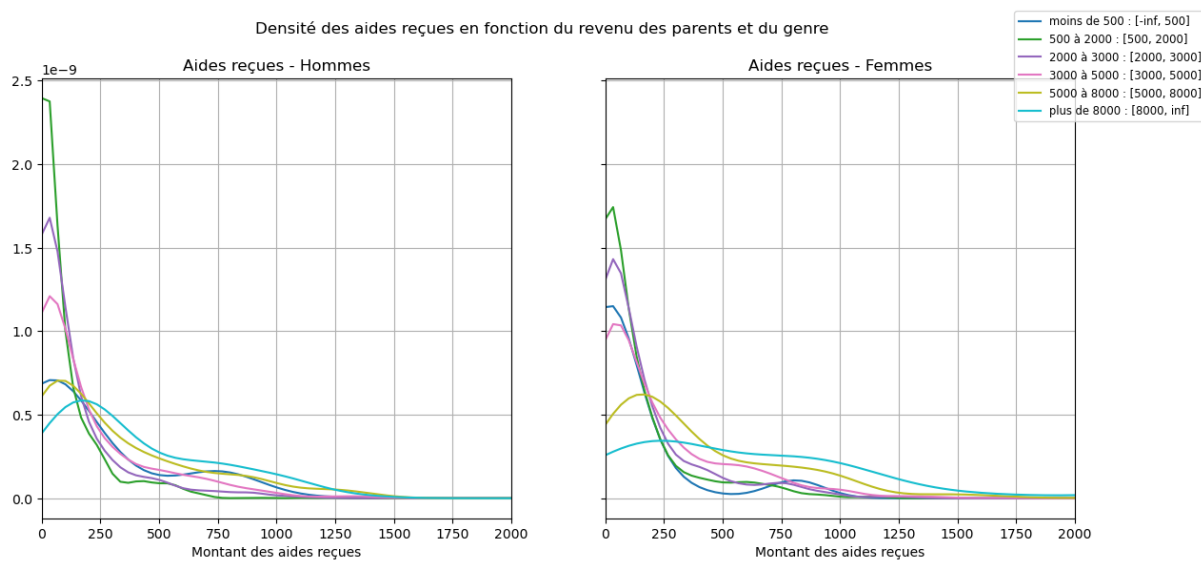


Figure 3: Distribution des aides en fonction du revenu et du sexe

SITPRINC	Hommes		Femmes	
	Moyenne	Écart-type	Moyenne	Écart-type
En emploi	56.83	118.18	48.20	109.07
Apprentissage ou stage rémunéré	131.02	206.97	145.16	226.71
En études ou en formation	338.41	311.48	305.82	307.32
Au chômage	88.79	143.01	83.93	128.98
Autre situation d'inactivité	87.62	146.41	112.60	168.55

Table 5: Statistiques par sexe et situation principale

en moyenne sauf quand il s'agit des jeunes en apprentissage ou en formation (différence dont il est dur de déterminer intuitivement l'origine) et pour les jeunes en autre situations d'inactivité. Peut-être est-ce lié au fait que cette catégorie inclut les femmes avec enfants (censées peser pour environ 68000 personnes contre même pas 4000 pour les hommes)?

## COHAB

COHAB	Hommes		Femmes	
	Moyenne	Écart-type	Moyenne	Écart-type
En cohabitation	106.92	133.26	96.79	127.85
En semi-cohabitation	452.27	336.93	425.87	339.69
Pas de cohabitation	183.32	308.28	154.43	287.00

Table 6: Statistiques par type de cohabitation et sexe

Ce qui est très étrange dans ce tableau, c'est que, pour chaque situation de cohabitation, les jeunes hommes sont plus aidés par leur parents que les jeunes femmes. Ainsi, on peut faire l'hypothèse que l'écart moyen est en faveur des jeunes femmes car elles sont surreprésentées dans une catégorie avec un fort niveau d'aide comme la catégorie "en semi-cohabitation". Pour vérifier cette intuition, il faut avoir recours à des regressions avec plusieurs variables explicatives dont COHAB.

## II Les femmes sont-elles toutes choses égales par ailleurs davantage aidées par leurs parents ?

Les individus de référence dans notre régression principale sont les jeunes hommes de 18 ans, encore en études, vivant avec leurs parents, dont chacun gagne en moyenne entre 500 et 2000€. Ainsi, nos interprétations porteront sur l'effet du fait de ne plus appartenir à cette catégorie de référence sur les aides parentales. Avant l'ajout des variables de contrôle, la régression OLS de l'aide parentale sur le sexe de l'enfant présentait un pouvoir explicatif très faible (avec un  $R^2$  ajusté d'environ 0,002). Après avoir intégré des variables de contrôle sélectionnées de manière cohérente avec les résultats du Lasso, le  $R^2$  ajusté atteint désormais 0,40, indiquant un modèle substantiellement plus explicatif.

Table 7: L'effet du sexe sur l'aide parentale contrôlé par l'âge, le revenu des parents, la situation professionnelle du jeune et la situation résidentielle vis-à-vis des parents.

	Tobit	OLS	Marge intensive	Effet marginal Logit)
intercept 1	314.3*** (1.3)	186.201*** (32.142)	47.339 (38.620)	
sexe féminin	20.1*** (0.2)	8.56 (5.568)	7.565 (6.829)	0.0301 *** (0.009)
âge	-11.5*** (6.6e-02)	-2.883 • (1.587)	3.400 • (1.920)	-0.0328*** (0.002)
revenu_r <500€	-109.4*** (1.3)	-57.846* (27.067)	-31.941 (43.907)	0.0074 (0.031)
revenu_r 2000-3000€	60.7*** (0.3)	33.993*** (7.432)	39.710*** (9.571)	0.0651*** (0.011)
revenu_r 3000-5000€	129.5*** (0.3)	91.563 *** (7.239)	98.921*** (9.108)	0.1039*** (0.011)
revenu_r 5000-8000€	249.8*** (0.4)	196.617*** (9.766)	199.194*** (11.484)	0.1855 *** (0.020)
revenu_r >8000€	333.8*** (0.7)	278.152*** (18.265)	271.510*** (20.055)	0.2237*** (0.050)
sitprinc emploi	-162.6*** (10.0)	-188.217*** (7.658)	-162.582*** (9.963)	-0.2686*** (0.010)
sitprinc apprentissage	-127.7*** (14.2)	-139.019*** (11.632)	-127.691*** (14.210)	-0.1910*** (0.016)
sitprinc chômage	-65.8*** (10.7)	-97.496*** (8.883)	-65.849*** (10.704)	-0.1609*** (0,014)
sitprinc autre	-70.6*** (17.7)	-118.516*** (13.969)	-70.632*** (17.665)	-0.1722*** (0.020)
semi-cohabitation	281.9*** (8.5)	251.705*** (7.621)	281.860*** (8.528)	0.0877*** (0.014)
non cohabitation	203.5*** (98.190)	7.448*** (0.1)	203.542*** (10.540)	-0.1112*** (0.009)
Observations	5197	5197	4113	5197
Multiple $R^2$	0.4033		0.4033	
Ajusted $R^2$	0.4014		0.4014	

## II.1 Influence des différentes variables de contrôle et conclusion sur l'impact du sexe sur l'aide des parents

### Le sexe

La régression OLS avec variables de contrôle indique que les jeunes filles reçoivent en moyenne toutes choses égales par ailleurs 8,56€ de plus que les garçons de la part de leurs parents, un écart qui n'est de plus pas statistiquement significatif. Cet écart était de 27,30€ dans la régression simple sans contrôle, ce qui suggère que l'essentiel de la différence brute s'explique par les variables ajoutées au modèle. Toutefois, une fois prise en compte la différence de probabilité de recevoir une aide entre jeunes femmes et jeunes hommes, les parents donnent en moyenne 20€ de plus aux filles. Cet écart, estimé à l'aide du modèle Tobit, est statistiquement significatif. L'examen séparé des marges intensive et extensive révèle que si l'effet de la marge intensive est à la fois faible en valeur et non significatif sur le plan statistique, la marge extensive, quant à elle, contribue davantage à l'écart observé entre jeunes hommes et jeunes femmes. Les filles ont 3% de chance de plus que les garçons de recevoir une quelconque aide financière de la part de leurs parents, chiffre plutôt faible, notamment par rapport à la régression sans variables de contrôle, mais tout de même significatif statistiquement. On en déduit que l'ajout des variables de contrôle réduit considérablement l'écart entre les hommes et les femmes en termes d'aide parentale, tant sur la marge intensive que sur la marge extensive. Cependant, si l'ajout des variables de contrôle rend la marge intensive non significative, la marge extensive demeure statistiquement significative. Une part de cette corrélation avec la marge extensive reste liée au genre, sans qu'aucune autre explication n'ait encore été avancée.

### L'âge

Lorsqu'on prend en compte non seulement la marge intensive, mais aussi la marge extensive — en traduisant en euros la variation de la probabilité de recevoir une aide —, on observe qu'une année supplémentaire d'âge fait perdre en moyenne 11,5€. Cet effet est significatif statistiquement. Toutefois, un examen plus approfondi révèle que seule la marge extensive est statistiquement significative, bien que son effet reste limité. Chaque année supplémentaire réduit de 3% la probabilité de recevoir une aide parentale. Cet effet reste dans tous les cas relativement faible et contribue peu à la réduction du coefficient associé au sexe. En effet, d'après la décomposition de Gelbach (voir le tableau IV.8), la prise en compte de l'âge ne réduit le coefficient OLS du sexe que de 0,045. Cela paraît cohérent car la répartition des jeunes filles et des jeunes hommes est équilibrée dans chaque classe d'âge. La limite de cette décomposition de Gelbach réside toutefois dans le fait qu'elle est calculée uniquement en fonction de la régression OLS, représentant mieux la marge intensive que la marge extensive.

### Le niveau de revenu des parents

Le niveau de revenu des parents a une influence très significative, tant sur le plan statistique que pratique, sur l'aide parentale. On observe que, toutes choses égales par ailleurs, plus les revenus des parents sont élevés, plus ils apportent une aide financière importante à leurs enfants. Par exemple, les jeunes issus de familles dont les parents gagnent chacun entre 3 000 € et 5 000 € ont un gain monétaire estimé, toutes choses égales par ailleurs, à 129,5 € par rapport à ceux dont les revenus parentaux se situent entre 500 € et 2 000 €. L'effet marginal estimé pour un passage de la tranche de revenus parentaux 500–2 000 € à plus de 8 000 € par parent est 2,43 fois plus élevé sur la marge intensive et 5,83 fois

plus élevé sur la marge extensive, comparativement à celui observé pour une transition vers la tranche 2 000–3 000 €. Le montant de l'aide parentale apportée au jeune croît avec le montant du revenu des parents à la fois sur les marges intensive et extensive, mais la sensibilité de la marge intensive au revenu est nettement plus prononcée.

La décomposition de Gelbach (voir tableau IV.8) met en évidence que l'ajout des tranches de revenu des parents comme variables de contrôle entraîne en moyenne une faible augmentation, de l'ordre de 0,32, du coefficient associé au sexe. Comme pour l'âge, cet impact très limité est cohérent, puisque les jeunes filles sont aussi nombreuses en pratique dans chacune des catégories. Par ailleurs, nous parlons de moyenne car chaque tranche de revenu est assigné à un coefficient différent (bien qu'il soit tous proches en valeur absolue et presque tous négatifs). L'importance relative des coefficients permet d'estimer des catégories sociales dans lesquelles les jeunes filles sont plus ou moins favorisées par rapport aux jeunes garçons en termes d'aide monétaire parentale. C'est parmi les individus dont les parents gagnent chacun entre 5 000 € et 8 000 € que l'écart selon le sexe est le moins marqué. À l'inverse, les parents gagnant chacun plus de 8 000 € sont ceux qui aident le plus leurs enfants avec une différenciation marquée entre filles et garçons. D'une part, les familles dont les parents gagnent plus de 8000€ chacun disposent de ressources financières importantes. Elles peuvent donc se permettre des stratégies d'allocation différenciée des aides selon le sexe, par exemple en favorisant davantage les filles. À l'inverse, les familles avec des revenus entre 5000 et 8000€ disposent certes de ressources très importantes, mais peut-être pas assez pour permettre à leurs jeunes enfants de vivre très confortablement tout en donnant plus à certains selon des critères corrélés avec le genre.

Cette interprétation demeure limitée, notamment en raison des contradictions qu'elle présente avec les observations issues de notre jeu de données. En effet, dans les catégories où les parents gagnent chacun moins de 5 000 €, l'écart genré est plus important que dans la tranche 5 000–8 000 €, bien que cette différence reste relativement faible. La différence pourrait donc ne pas s'expliquer entièrement par le niveau de revenu lui-même, mais également par la catégorie socio-professionnelle, qui est corrélée au revenu et dont l'effet pourrait ainsi être partiellement absorbé par le niveau de revenu dans notre régression. Par exemple, les professeurs d'université pourraient avoir tendance à soutenir davantage financièrement leurs enfants pour leurs études, alors que des parents aux revenus similaires mais issus de la haute sphère économique plutôt qu'intellectuelle pourraient préférer que leurs enfants contribuent eux-mêmes au financement de leurs études. Il pourrait par conséquent être pertinent d'inclure la catégorie socio-professionnelle comme variable explicative, tout en restant très prudent dans l'interprétation des résultats, compte tenu de sa corrélation avec le niveau de revenu, afin de distinguer clairement les effets respectifs de chaque variable.

## La situation professionnelle

Les jeunes ayant déjà intégré le marché du travail sont beaucoup moins aidés par leurs parents que les étudiants. Un jeune entre 18 et 24 ans occupant un emploi reçoit en moyenne, *ceteris paribus* et en prenant en compte le gain monétaire procuré par la marge extensive ainsi que celui résultant de la marge intensive, 162,6€ de moins qu'un étudiant. Cet effet est significatif statistiquement et ce n'est pas étonnant.

Ce qui surprend en revanche, c'est qu'un jeune en apprentissage reçoive en moyenne moins d'aide qu'un jeune au chômage. Un jeune au chômage reçoit 1,48 fois plus d'aide parentale sur la marge intensive et est 1,16 fois plus aidé sur la marge extensive qu'un jeune en alternance. L'écart est ainsi plus marqué sur la marge intensive que sur la marge extensive. S'il nous faudrait avoir accès à plus de données pour expliquer cet écart, on peut



avancer plusieurs hypothèses. D'une part, l'apprentissage apparaît comme une situation plus stable et pérenne que le chômage. De ce fait, les parents de jeunes en apprentissage sont peut-être davantage rassurés et considèrent leurs enfants comme plus autonomes, même si la rémunération minimale en apprentissage reste inférieure à l'allocation minimale perçue au chômage. D'autre part, il serait nécessaire de disposer d'une variable indiquant la durée du chômage, car des individus ne percevant plus d'allocations, ou très peu en raison d'une période de chômage prolongée, pourraient recevoir un soutien parental nettement plus important et ainsi biaiser les résultats sur l'ensemble de l'échantillon.

Enfin, il serait nécessaire de disposer de davantage d'informations sur les situations les plus fréquentes regroupées dans la catégorie "autre", si l'on souhaite en tirer une interprétation pertinente.

L'introduction de la variable relative à la situation professionnelle contribue fortement à la diminution du coefficient associé au sexe dans notre régression linéaire. Parmi les modalités de la situation professionnelle, c'est le fait d'être en emploi qui explique le plus fortement l'écart d'aide parentale observé entre les hommes et les femmes. On en déduit que les femmes sont sous-représentées parmi les jeunes en emploi, mais aussi plus généralement surreprésentées parmi les étudiants en formation initiale. Étant surreprésentées dans une situation où les aides parentales sont à la fois plus fréquentes et plus importantes, il est logique qu'elles bénéficient finalement d'un soutien parental plus élevé que les jeunes garçons, même si cette variable n'explique pas totalement l'écart initial observé entre les jeunes hommes et les jeunes femmes.

## **La cohabitation ou non avec les parents**

Tout comme le revenu des parents ou encore la situation professionnelle du jeune adulte, la cohabitation ou non du jeune avec ses parents est très corrélée avec le montant d'aide de ces derniers que le jeune reçoit. Étonnamment, un jeune en semi-cohabitation reçoit en moyenne plus d'aides de ses parents *ceteris paribus* qu'un jeune n'habitant plus du tout avec ses parents. Pourtant, on pourrait légitimement considérer que ce dernier a des besoins financiers plus importants.

Nous avons formulé deux principales hypothèses susceptibles d'expliquer les résultats observés, bien qu'une analyse plus approfondie du jeu de données soit nécessaire pour les confirmer. D'un côté, il est possible que les jeunes en semi-cohabitation entretiennent de meilleures relations avec leurs parents que ceux en situation de non-cohabitation totale. En effet, à cet âge, il est rare d'être totalement autonome. Le fait de ne plus vivre du tout chez ses parents peut traduire un éloignement volontaire — ou contraint — résultant de relations familiales difficiles. Le fait que les jeunes en semi-cohabitation soient aidés 11% moins souvent que ceux vivant encore chez leurs parents, tout en recevant en moyenne 203,54€ de plus sur la marge intensive, semble conforter ce postulat. Parmi ces jeunes, certains pourraient être en mauvais termes avec leurs parents et ne percevoir que peu ou pas d'aide, ce qui pourrait biaiser les coefficients estimés. D'un autre côté, il est également possible que les jeunes en situation de non-cohabitation présentent des caractéristiques différentes — par exemple, vivre en couple —, ce qui pourrait expliquer un soutien parental plus limité. Toutefois, l'une des principales caractéristiques susceptibles de les différencier — à savoir le fait d'être en emploi — a déjà été prise en compte dans notre analyse. De prime à bord, la première hypothèse nous paraît donc à privilégier.

À l'instar de la situation professionnelle, le fait de prendre en compte la cohabitation avec les parents comme variable de contrôle contribue considérablement à la réduction du coefficient associé au sexe. On en conclut que les jeunes filles vivent plus fréquemment hors du domicile parental, et surtout en situation de non-cohabitation totale, même après

contrôle de la situation professionnelle. Cela suggère que cette tendance ne s'explique pas uniquement par leur surreprésentation dans le milieu des études, qui impose souvent aux jeunes adultes de s'éloigner, au moins temporairement, du foyer familial. Cet écart pourrait justement être lié au sexe du jeune adulte et aux mécanismes de socialisation genrée. Par conséquent, l'inclusion de la variable de cohabitation pourrait partiellement masquer l'effet direct du sexe sur l'aide parentale. Malgré tout, il est probable que d'autres variables, corrélées au sexe sans en être la conséquence directe, expliquent également cet écart, telles que la situation matrimoniale. On peut en effet supposer que les femmes se retrouvent plus jeunes dans des relations de couple stables et durables, notamment en raison du fait que, dans les relations hétérosexuelles, les hommes sont en moyenne plus âgés que leurs partenaires féminines.

## II.2 Limites du modèle Tobit I pour mesurer les marges intensives et extensives

Dans notre tableau comparant les différents coefficients de régression, on constate que les coefficients estimant séparément la marge extensive et la marge intensive évoluent différemment. De ce fait, le modèle Tobit I n'est pas forcément adapté pour étudier simultanément la marge intensive et la marge extensive, puisque l'une de ses hypothèses fondamentales est que ces deux marges sont déterminées de manière identique, par les mêmes variables, et évoluent de façon similaire. Or, dans notre régression, après ajout des variables de contrôle sélectionnées, la marge extensive reste statistiquement significative pour expliquer le coefficient associé au sexe, alors que la marge intensive ne l'est plus. Il existe donc probablement des variables qui influencent davantage la décision d'aider ou non son enfant, mais qui ont un impact bien moindre sur le montant final accordé. La décision de donner ou non ne se confond pas avec celle de déterminer le montant à accorder à son enfant. Par conséquent, il pourrait être plus pertinent d'adopter un modèle Tobit II pour ces données, ce modèle traitant distinctement la marge extensive et la marge intensive.

## III Un écart sexué qui dépend du type d'aide

### III.1 Types d'aides parentales : niveau et fréquence

In fine, même si l'écart du montant total des aides parentales entre les jeunes femmes et les jeunes hommes est très faible, voire statistiquement non significatif, cela ne signifie pas pour autant qu'aucune différence n'existe. En effet, aidetot étant une variable agrégée, il est tout à fait possible que de fortes différences entre les sexes existent pour chaque type d'aide, mais qu'elles soient masquées par l'agrégation de l'ensemble des aides. C'est pourquoi nous avons souhaité examiner plus en détail les différences éventuelles dans les types d'aides reçues selon le sexe des jeunes.

	aide_logement	aide_loisirs	aide_santé	aide_scolarité	aide_communication	aide_transports	aide_vêtements	aide_voyages
Femme	41.929	73.253	11.898	26.967	51.571	56.986	50.746	19.147
Homme	33.195	79.274	10.363	23.077	50.46	54.075	50.929	16.672

Figure 4: Tableau représentant les pourcentages de femmes et d'hommes recevant chaque type d'aide

Le tableau ci-dessus, représentant le pourcentage d'hommes et de femmes recevant chaque type d'aide, nous montre qu'effectivement les jeunes hommes et les jeunes femmes ne reçoivent pas les mêmes aides de leurs parents. Nous remarquons par exemple, que 41,9% des femmes reçoivent une aide financière pour le logement de la part de leurs parents, contre seulement 33,2% des hommes. Ainsi, les jeunes femmes semblent à première vue, être légèrement favorisées par leurs parents en ce qui concerne le logement, la scolarité et les voyages<sup>1</sup>, tandis que les jeunes hommes le sont en ce qui concerne les loisirs. Cela concerne l'aide à la marge extensive des parents envers leurs enfants, c'est-à-dire que ce tableau ne prend en compte que la fréquence à laquelle l'aide est perçue dans chaque groupe (le nombre de femmes et d'hommes qui reçoivent un certain type d'aide).

	montant_aide_logement	montant_aide_loisirs	montant_aide_santé	montant_aide_scolarité	montant_aide_communication	montant_aide_transports	montant_aide_vêtements	montant_aide_voyages
Moyenne femmes	340.381	57.144	22.898	1375.65	28.044	72.988	293.636	856.373
Médiane femmes	295	30	20	400	23.5	54.583	200	400
Moyenne hommes	333.081	54.185	21.823	1604.348	27.368	66.012	289.189	756.952
Médiane hommes	210	30	20	405	20	50	200	400

Figure 5: Tableau représentant les médianes et moyennes genrées des montant reçus par type d'aide

Il semble également intéressant d'étudier l'aide à la marge intensive. Ainsi, dans le tableau ci-dessus nous pouvons lire les moyennes et médianes genrées des montants reçus par les enfants de la part de leurs parents selon le type d'aide<sup>2</sup>. Ces moyennes et médianes ont été calculées sur le sous-ensemble des jeunes ayant renseigné un montant non-nul pour le type d'aide en question.

Ici, nous remarquons que la médiane est plus élevée chez les femmes recevant une aide au logement que chez les hommes recevant cette même aide. De même, pour l'aide pour les voyages, l'aide moyenne perçue par les jeunes femmes est supérieure de 100€ à celle perçue par les jeunes hommes. Toutefois, la médiane est la même ce qui suggère qu'il y a plus de valeurs extrêmes pour les montants reçus par les jeunes femmes que par les jeunes hommes. En revanche, contrairement à ce que suggérerait le tableau sur l'aide à la marge extensive, le niveau d'aide que les jeunes reçoivent de la part de leurs parents pour leurs dépenses de loisirs semble assez équivalent selon que le jeune est un homme ou une femme.

Ces statistiques descriptives nous invitent donc à poursuivre notre étude de l'apport parental aux jeunes selon leur genre et la catégorie d'aide sélectionnée.

## III.2 Inégalités d'accès à l'aide parentale selon le sexe et le type d'aide

### III.2.1 Inégalité genrée d'accès à l'aide parentale pour le logement

Comme nous l'avons fait remarquer dans la partie précédente, il existe des différences dans les types d'aides reçus par les jeunes femmes et les jeunes hommes de la part de leurs parents. Notamment, nous avons pu mettre en lumière le fait que 41,9% des jeunes femmes reçoivent une aide financière de la part de leurs parents pour se loger contre seulement 33,2% des jeunes hommes. Aussi, nous avons effectué plusieurs régressions logistiques permettant d'approximer l'effet causal du fait d'être une femme sur le fait de recevoir une aide au logement de la part de ses parents.

<sup>1</sup>Voir annexes pour la construction des variables binaires d'aide par catégorie.

<sup>2</sup>Voir annexes pour la construction des variables de montant en fonction des types d'aides.

Table 8: Effets marginaux logement - Modèle sans l'aide du partenaire

Variable	Effet marginal	p-value
agej_scaled	0.0177*	0.0292
apl_binaire	0.1528***	1.94e-17
cohab_binaire	-0.3773***	3.539e-304
femme	0.0287*	0.04147
revenu_r_scaled	0.0567***	5.084e-12
sitprinc_Apprentissage	-0.0925***	0.0003301
sitprinc_Autre	-0.1222**	0.009993
sitprinc_Chomage	-0.0798**	0.004259
sitprinc_Emploi	-0.045*	0.01728

*Note:* \*p<0.05; \*\*p<0.01; \*\*\*p<0.001

Pour cela, nous avons ajouté plusieurs variables de contrôle : tout d'abord, nous avons ajouté les variables de contrôle utilisées dans toutes les autres régressions (comprenant le revenu des parents, l'âge du jeune, le fait qu'il vive ou non chez ses parents<sup>3</sup> et sa situation professionnelle<sup>4</sup>) ; nous avons également ajouté le fait de toucher des APL ou non, et le fait d'être aidé par son conjoint ou partenaire<sup>5</sup> comme variables de contrôle. Nous avons choisi de calculer les effets marginaux moyens sur le modèle n'incluant pas l'aide du partenaire comme variable de contrôle. Nous avons fait ce choix car il nous a semblé, intuitivement mais également au vu des résultats, que cette variable pouvait introduire un biais de simultanéité dans notre modèle. En effet, il n'est pas possible de savoir si l'aide du partenaire vient combler une absence d'aide parentale ou si au contraire les parents décident de ne plus aider leur enfant car le partenaire le fait déjà.

Dans cette étude, nous utilisons également une variable continue normalisée de revenu comme variable de contrôle plutôt qu'une variable par tranches, car la variable expliquée (les aides par catégorie) concerne un nombre limité d'individus dans mon échantillon. En effet, moins de personnes reçoivent cette aide, ce qui crée une distribution plus concentrée et moins homogène parmi les bénéficiaires. Une variable continue permet ainsi de mieux capturer les variations subtiles de revenu, notamment près des seuils critiques.

Ici, notre étude montre que toute choses égales par ailleurs le fait d'être une femme augmente en moyenne la probabilité de recevoir une aide au logement de la part de ses parents de 2,85 points de pourcentage. Notons que dans le modèle logistique pour lequel nous avons calculé ce coefficient<sup>6</sup>, la p-value du coefficient associé à la variable explicative **femme** est inférieure à 0,05 mais supérieure à 0,01. Cela s'explique très probablement par un manque d'observations croisant toutes les variables de contrôle. Toutefois, les résultats restent solide et il semble donc y avoir un réel effet du sexe sur le fait de recevoir une aide financière de la part de ses parents pour se loger.

<sup>3</sup>Voir annexes pour la construction de la variable.

<sup>4</sup>Voir annexes pour la construction de la variable.

<sup>5</sup>Voir annexes pour la construction de la variable.

<sup>6</sup>La troisième régression du tableau, contenant toutes les variables de contrôle sauf **aide.partenaire.logement**.

### III.2.2 Inégalité genrée d'accès à l'aide parentale pour les loisirs

Nous avons également pu noter que les femmes semblaient être défavorisées par leurs parents en ce qu'il est de recevoir une aide pour les loisirs. En effet, alors que 79,3% des hommes perçoivent une aide pour leurs dépenses de loisirs de la part de leurs parents, seules 73,6% des femmes en perçoivent une. Aussi, nous avons décidé d'effectuer plusieurs régressions logistiques pour tenter d'identifier l'effet causal du sexe sur le fait de recevoir ou non une aide pour les dépenses de loisirs.

Table 9: Effets marginaux loisirs - Modèle sans l'aide du partenaire

Variable	Effet marginal	p-value
agej_scaled	0.0135	0.1028
cohab_binaire	0.0166	0.2578
femme	-0.0566***	3.732e-05
revenu_r_scaled	0.0802***	6.852e-09
sitprinc_Apprentissage	-0.032	0.2824
sitprinc_Autre	-0.1807***	4.448e-09
sitprinc_Chomage	-0.1847***	1.737e-20
sitprinc_Emploi	-0.0564**	0.004165

*Note:* \*p<0.05; \*\*p<0.01; \*\*\*p<0.001

Nous avons ainsi effectué trois régressions logistiques : Une première n'incluant aucune variable de contrôle ; Une deuxième incluant les variables de contrôle usuelle ; Et une dernière incluant en plus la variable d'aide du partenaire qui nous permet de savoir si le jeune est aidé ou non par son partenaire. Comme précédemment, nous avons décidé de calculer les coefficients marginaux moyens sur le deuxième modèle puisqu'il semble encore probable que la variable `aide_partenaire_loisirs` introduise un biais de simultanéité dans notre modèle. Nous pouvons également noter que le résultat obtenu pour le coefficient associé à la variable explicative `femme` est solide comme le confirme la p-value inférieure à 0,001.

Ainsi, notre étude nous permet de conclure que toutes choses égales par ailleurs, le fait d'être une femme fait baisser la probabilité de recevoir une aide pour les loisirs de la part de ses parents d'en moyenne 5,66 points de pourcentage. Il y a donc un effet négatif du fait d'être une femme sur le fait de recevoir une aide de la part de ses parents pour financer ses dépenses de loisirs.

Ces régressions montrent que, selon les catégories d'aides, il peut être avantageux ou désavantageux d'être une femme. Nous remarquons ainsi que les jeunes femmes sont souvent aidées en matière de logement, tandis que les jeunes hommes sont plus souvent aidés en ce qui concerne les loisirs. Cela semble particulièrement intéressant à relever : les jeunes hommes bénéficieraient donc plus facilement d'un accès aux loisirs grâce à l'aide de leurs parents, tandis que les jeunes femmes auraient un meilleur accès au logement, et plus précisément à un logement personnel.

Nous avons également réalisé des régressions pour identifier l'effet du sexe sur la probabilité de recevoir une aide pour la santé, la scolarité, la communication, les transports, les vêtements et les voyages. Cela nous a permis de mettre en évidence le fait que les femmes sont également favorisées en ce qui concerne les transports. Pour les autres catégories

d'aides cependant, les coefficients estimés restent peu significatifs. Toutefois, l'effet estimé est positif pour toutes les autres catégories, à l'exception de la catégorie vêtements pour laquelle l'effet est très faiblement négatif pour une p-value très élevée<sup>7</sup>. Dans l'ensemble, ces résultats confirment ce que nous avons identifié dans les parties précédentes : le fait d'être une femme augmente la probabilité de recevoir une aide financière de la part de ses parents.

### III.3 Ecarts de montant des aides reçues entre les hommes et les femmes

Nous avons vu qu'il existait des inégalités d'accès aux aides parentales selon le type d'aide et le sexe du jeune concerné. Pour compléter cette étude des aides parentales à la marge extensive, il semble pertinent de mener une étude des aides parentales à la marge intensive. Pour cela, nous avons créé des variables répertoriant les montants reçus pour chaque type d'aide et avons ensuite effectué des régressions pour tenter d'évaluer l'effet du sexe sur le montant reçu.

Trois choses sont à noter concernant cette étude : tout d'abord, nous avons mené l'étude uniquement sur les jeunes recevant une aide en excluant ceux qui déclaraient ne pas recevoir d'aide ainsi que ceux n'ayant renseigné aucun montant pour l'aide en question ; ensuite, nous avons remarqué que pour tous les types d'aides, aucun des jeunes n'ayant affirmé ne pas recevoir un certain type d'aide n'a renseigné un montant supérieur à zéro pour cette même aide, ce qui est positif. Toutefois, pour certaines catégories d'aides, une partie, parfois conséquente, des jeunes ayant affirmé recevoir l'aide, n'a pas indiqué de montant. Nous avons considéré ce phénomène d'attrition en créant des tableaux montrant la distribution des observations entre les différentes variables de contrôle, pour les jeunes déclarant une aide mais aucun montant, ainsi que pour les jeunes déclarant une aide et un montant positif. Cela nous a permis de conclure que cette attrition était effectivement sélective, ce qu'il faut garder à l'esprit lorsque l'on étudie les résultats de ces régressions<sup>8</sup>. Enfin, nous avons décidé de ne pas réaliser de régressions sur les variables de montant pour lesquelles nous disposions de moins de 500 observations supérieures à zéro.

---

<sup>7</sup>Voir annexes pour le reste des régressions logistiques.

<sup>8</sup>Voir annexes pour les tableaux et les détails concernant l'attrition.

### III.3.1 Ecarts de montant de l'aide reçue pour le logement entre les hommes et les femmes

	<i>Dependent variable:</i>		
	montant_aide_logement		
	Sans contrôle (1)	Contrôles usuels (2)	Avec APL (3)
femme	5.434 (0.725)	11.394 (0.413749)	9.773 (0.485576)
revenu_r		0.021*** (4.422e-05)	0.022*** (3.023e-05)
agej		2.952 (0.527960)	-0.207 (0.968596)
sitprinc_Emploi		-80.722 (0.127489)	-76.478 (0.139454)
sitprinc_Apprentissage		-105.720** (0.003082)	-105.090** (0.003236)
sitprinc_Chomage		65.537 (0.372737)	71.037 (0.353419)
sitprinc_Autre		32.392 (0.549902)	33.640 (0.542672)
apl_binaire			32.934* (0.035628)
Constant	331.487*** (<2e-16)	188.046* (0.048148)	230.188* (0.025527)
Observations	1,166	1,166	1,166
R <sup>2</sup>	0.0002	0.121	0.127
Residual Std. Error	4,213.217 (df = 1164)	3,961.848 (df = 1158)	3,947.818 (df = 1157)
Adjusted R <sup>2</sup>	-0.001	0.115	0.121
Residual Std. Error	4,213.217 (df = 1164)	3,961.848 (df = 1158)	3,947.818 (df = 1157)
F Statistic	0.268 (df = 1; 1164)	22.671*** (df = 7; 1158)	21.134*** (df = 8; 1157)

Note:

\*p<0.05; \*\*p<0.01; \*\*\*p<0.001

Concernant l'aide pour le logement, notre étude ne permet pas de conclure avec significativité sur l'effet du sexe sur le montant reçu par les jeunes percevant l'aide. En effet, les coefficients associés à la variable explicative **femme** ont une p-value supérieure à 0,05 pour les trois régressions effectuées. Cela s'explique par le faible nombre d'observations : 1 166 ici. Toutefois, notons que les coefficients sont élevés, à hauteur d'environ 9,77. Cela signifie que, selon cette estimation, toutes choses égales par ailleurs, le fait d'être une femme augmenterait de 9,77€ l'aide au logement mensuelle reçue par un jeune. Cela n'est pas négligeable, il serait donc intéressant de pouvoir effectuer des régressions sur de plus grands échantillons afin de vérifier si la tendance se maintient.

### III.3.2 Ecarts de montant de l'aide reçue pour les loisirs entre les hommes et les femmes

Nous nous intéressons maintenant à l'aide des parents pour voyager. En effet, nous avons remarqué que les femmes recevant une aide de la part de leurs parents recevaient en moyenne 100 € de plus que les hommes dans la même situation. Aussi, pour vérifier l'hypothèse selon laquelle les femmes seraient avantagées par leurs parents dans les montants qu'elles reçoivent pour voyager, nous avons effectué deux régressions linéaires : Une

sans variable de contrôle et une avec les variables de contrôle usuelles. En ce qui concerne l'attrition, pour ce type d'aide, aucun phénomène d'attrition n'a été repéré, les jeunes ayant déclaré recevoir une aide ont déclaré un montant et les jeunes n'ayant pas déclaré d'aide n'ont pas déclaré de montant. Nous avons donc pu réaliser les régressions sur les 1 089 jeunes ayant déclaré une aide.

Table 10: Régressions linéaires pour l'aide concernant les voyages

	<i>Dependent variable:</i>	
	montant_aide_voyages	
	Sans contrôle (1)	Contrôles usuels (2)
femme	89.931 (0.2433)	100.714 (0.1665)
revenu_r		0.074*** (2.739e-05)
agej		38.819 (0.1019)
sitprinc_Emploi		-201.370 (0.0761)
sitprinc_Apprentissage		-505.100*** (3.899e-10)
sitprinc_Chomage		-43.847 (0.7581)
sitprinc_Autre		-189.170 (0.1972)
cohab_binaire		-102.128 (0.1886)
Constant	754.135*** ( $<2e-16$ )	-218.761 (0.6510)
Observations	1,040	1,040
R <sup>2</sup>	0.002	0.065
Adjusted R <sup>2</sup>	0.001	0.057
Residual Std. Error	30,474.880 (df = 1038)	29,604.520 (df = 1031)
F Statistic	2.083 (df = 1; 1038)	8.892*** (df = 8; 1031)

*Note:* \*p<0.05; \*\*p<0.01; \*\*\*p<0.001

Toutefois, cela reste peu et nos résultats le montrent. Très peu des coefficients estimés sont significatifs. Cela dit, notons que le coefficient associé à la variable **femme** est très élevé (supérieur à 100 € annuels). Aussi, nous pouvons supposer que le fait d'être une femme représente en effet un avantage dans le montant perçu de la part de ses parents pour voyager. Il serait alors intéressant de mener une étude sur un échantillon plus large pour en déduire des conclusions plus solides.

L'étude des autres types d'aide, bien que peu significative en raison du faible nombre d'observations, tend à montrer que les femmes sont avantagées en ce qui concerne les loisirs (contrairement à ce qu'indique l'étude de l'aide à la marge extensive), les transports, les vêtements ; et désavantagées en ce qui concerne la scolarité<sup>9</sup>.

<sup>9</sup>Voir annexes pour le reste des régressions linéaires par catégorie d'aide.



Ainsi, notre étude montre qu'il existe bel et bien une aide parentale genrée : les jeunes femmes et les jeunes hommes ne sont pas aidés au même niveau mais surtout ils ne sont pas aidés de la même manière et pas pour les mêmes choses. Cela est particulièrement intéressant à relever puisqu'au-delà du fait de percevoir ou non une aide parentale, les hommes et les femmes ne sont donc pas encouragés par leurs parents dans les mêmes activités ni les mêmes domaines de leur vie personnelle ou professionnelle.

## Conclusion

Notre étude de la base de données ENRJ nous a permis d'arriver à plusieurs résultats au sujet de l'aide que les parents fournissent à leurs enfants. Ces résultats concernent à la fois l'ampleur des différences genrées dans ces aides, ainsi que les facteurs, au-delà du sexe, qui influencent la décision de fournir une aide, son montant, et qui expliquent en partie l'effet du sexe lui-même. Ce qu'il faut retenir, c'est que, sans prise en compte des variables de contrôle, les jeunes femmes semblent bénéficier à la fois d'un montant d'aide plus élevé et d'une probabilité plus grande de recevoir une aide.

Cependant, après avoir introduit comme variables de contrôle l'âge, le revenu des parents, la situation professionnelle du jeune et la cohabitation avec les parents, on observe que le sexe a un impact faible sur le montant total des aides parentales, tant sur la marge intensive que sur la marge extensive. Plus précisément l'effet du sexe sur la marge intensive, en plus d'être faible, n'est pas statistiquement significatif, alors que son effet sur la marge extensive l'est (bien que faible).

Comment alors expliquer la réduction de cet écart qui était pourtant singulier ? Il ne semble pas que l'âge ou le revenu des parents expliquent cette réduction de l'écart en fonction du sexe des aides parentales, puisque les jeunes hommes et femmes sont répartis de manière équilibrée selon ces catégories. En revanche, la situation professionnelle du jeune et son statut de cohabitation avec ses parents expliquent, eux, en grande partie cet écart. L'ajout de ces variables en tant que contrôles permet donc de réduire drastiquement le coefficient associé au sexe. Cela s'explique par le fait que les jeunes femmes sont surreprésentées dans les catégories de ces deux variables pour lesquelles l'aide parentale est la plus élevée (semi-cohabitantes ou non-cohabitantes et en études), contrairement aux jeunes hommes qui restent plus longtemps chez leurs parents et entrent plus tôt sur le marché du travail.

Cependant, cet écart global masque peut-être une réalité beaucoup plus contrastée. C'est pourquoi nous avons choisi d'étudier séparément les différents types d'aide, afin de pouvoir tirer des conclusions plus précises de notre analyse. En effet, il se pourrait qu'un écart significatif et incompressible d'aide parentale selon le sexe existe, mais que cet écart soit peu visible à l'échelle d'une variable agrégée et ne se manifeste réellement qu'en examinant les différences d'aide secteur par secteur.

Bien que les résultats restent peu significatifs, l'analyse par type d'aide suggère que les jeunes femmes bénéficient davantage d'aides parentales que les jeunes hommes dans le cadre de leurs loisirs (contrairement à ce que montre l'étude sur la marge extensive), de l'usage de transports ainsi que de l'achat de vêtements, tandis qu'elles sont moins favorisées en matière d'aide pour suivre leur scolarité<sup>10</sup>.

Par conséquent, il existe bel et bien une aide parentale genrée : même si la différence en faveur des jeunes femmes semble faible, notamment une fois prises en compte les variables de contrôle dans la régression de l'aide totale perçue selon le sexe, les jeunes hommes et femmes ne bénéficient pas du même type d'aide ni de la même manière. Cela est particulièrement intéressant à relever puisqu'au-delà du fait de percevoir ou non une aide parentale, les hommes et les femmes ne sont donc pas encouragés par leurs parents dans les mêmes activités ni les mêmes domaines de leur vie personnelle ou professionnelle.

---

<sup>10</sup>Voir annexes pour les autres régressions linéaires par catégorie d'aide.

## IV Annexe

### IV.1 Opérations sur la base de données

L'enquête ENRJ a conduit à la création de trois bases de données. Pour mener nos analyses, nous avons mobilisé deux d'entre elles : la base issue des questionnaires adressés aux jeunes et celle provenant des questionnaires adressés à leurs parents. Afin de croiser les informations issues de ces deux sources, nous avons procédé à leur fusion. L'identifiant commun aux deux bases, permettant d'apparier chaque jeune à ses parents, est l'identifiant `idmen`. Chaque jeune adulte a un `idmen` différent. En revanche, chaque parent n'a pas forcément un `idmen` unique. En effet, il arrive que les deux parents d'un même jeune répondent au questionnaire, ce qui entraîne une double occurrence de l'identifiant `idmen` dans la base correspondante. Nous avons déterminé qu'il y avait 709 `idmen` sur 5776 qui apparaissaient plusieurs fois dans la base, nombre petit mais non négligeable pour autant. Il n'était pas possible de procéder à une simple jointure ou juxtaposition des tables. Par ailleurs, exclure les identifiants `idmen` concernés aurait pu introduire un biais de sélection, dans la mesure où les jeunes dont les deux parents ont répondu peuvent partager certaines caractéristiques spécifiques. Pour permettre la fusion des bases, nous avons d'abord modifié la base de données des parents afin de n'avoir qu'une seule ligne par identifiant `idmen`. Pour cela, chaque colonne a été dupliquée de manière à pouvoir accueillir les réponses d'un éventuel second parent répondant. Une fois cette base restructurée, nous avons pu procéder à la jointure avec la base des jeunes.

```
1 # creation d'un dataframe avec une nouvelle colonne indiquant un
2 # numero de parent pour chaque idmen
3 parents_num <- data_parents %>%
4   group_by(idmen) %>%
5   mutate(num_parent = row_number()) %>%
6   ungroup()
7
8 cols_vals <- dplyr::select(parents_num, -idmen, -num_parent)
9 %>% colnames()
10 # je rajoute dplyr. Sinon le code ne pointait pas vers la bonne
11 #fonction
12
13 # concatenation des variables informant sur les caractéristiques
14 # des deux parents, afin d'avoir une ligne par enfant
15 # et non pas par parent
16 data_parents_finale <- parents_num %>%
17   pivot_wider(
18     id_cols = idmen,
19     names_from = num_parent,
20     values_from = all_of(cols_vals),
21     names_sep = ""
22   )
23
24 # merge des bases
25 data <- data_jeunes %>%
26   left_join(data_parents_finale, by = "idmen")
```

Listing 1: Fusion des bases des questionnaires pour les jeunes et pour les parents

## IV.2 Questionnements méthodologiques sur le LASSO

Pour effectuer le LASSO, nous avons dû réduire le nombre de variables explicatives possibles. En effet, le nombre de variables étaient très conséquents, beaucoup d'entre elles présentant un problème d'endogénéité avec la variable que nous voulions expliquer, d'autres étaient endogènes entre elles. En réalité, beaucoup de variables étaient très proches ou dérivées d'autres variables. AIDETOT est la résultante d'une transformation de nombreuses autres variables d'aides plus précises (cela est précisé dans le dictionnaire des variables). Les variables qui composent AIDETOT sont elle-même dérivées d'autres variables et ainsi de suite. Par conséquent, il était nécessaire d'exclure un certain nombre de variables. Or, le nombre de variables d'aides étant très élevé, pour le LASSO sur l'aide totale mensuelle perçue par les jeunes, nous avons plutôt décidé de sélectionner une trentaine de variables explicatives qui nous semblaient pertinentes et qui ne se recoupaient pas trop. Pour le LASSO sur la satisfaction des jeunes par rapport à l'aide perçue, nous avons sélectionné par un processus éliminatif les variables que nous incluions ou non. Concrètement, nous avons pris les 30 premières variables avec les plus gros coefficients, puis nous avons éliminé les variables qui se recoupaient, s'il n'y avait pas de changement sur le  $R^2$  du LASSO.

## IV.3 Construction de tranches de revenus

Il a d'abord été nécessaire de déterminer une méthode de calcul du revenu parental une fois les bases fusionnées, afin d'obtenir une variable unique agrégée, en lieu et place des deux variables distinctes correspondant à chacun des parents. Il n'était pas envisageable de simplement additionner les revenus des deux parents dans les cas où chacun avait répondu au questionnaire. En effet, additionner les revenus des deux parents lorsque chacun a répondu au questionnaire conduirait mécaniquement à des niveaux de revenu parental plus élevés pour ces jeunes, par rapport à ceux dont un seul parent a répondu. Une telle approche introduirait un biais important dans l'analyse. Nous avons donc choisi de calculer la moyenne des revenus parentaux lorsque les deux sont renseignés, afin de garantir une comparaison cohérente entre les individus. Une des limites de notre présente analyse est l'absence de distinction entre familles monoparentales et autres, ce qui peut potentiellement biaiser les résultats.

```
1 data <- data %>%
2   mutate(revenu_r = case_when(
3     is.na(revenu_r2) ~ revenu_r1,
4     is.na(revenu_r1) ~ revenu_r2,
5     !is.na(revenu_r1)&!is.na(revenu_r2) ~ (revenu_r1 + revenu_r2)/2,
6     TRUE ~ NA_real_
7     # Cela permet de gerer les valeurs NA ou non definies
8   ))
```

Listing 2: Création de la variable indiquant le revenu des parents

À partir de cette nouvelle variable `revenu_r`, nous avons construit des catégories de tranches de revenu. Nous avons d'abord recréé une variable reprenant les mêmes tranches de revenu que celles présentes dans le jeu de données initial. Cependant, compte tenu du nombre important de catégories, nous avons regroupé certaines tranches afin de rendre les tableaux de régression plus lisibles et interprétables.

Pour guider ce choix, nous avons effectué une régression du montant total des aides parentales sur les tranches de revenu déjà existantes afin d'identifier et regrouper celles

OLS	
(Intercept)	216.32*** (8.35)
moins_de_500	-175.33*** (15.20)
de_500_1000	-140.81*** (13.58)
de_1000_1250	-123.09*** (16.43)
de_1250_1500	-139.91*** (10.67)
de_1500_2000	-112.84*** (11.00)
de_2000_2500	-89.04*** (11.95)
de_2500_3000	-67.45*** (13.11)
de_5000_8000	138.05*** (19.72)
plus_de_8000	233.28*** (47.81)
R <sup>2</sup>	0.13
Adj. R <sup>2</sup>	0.13
Num. obs.	5197

\*\*\*  $p < 0.001$ ; \*\*  $p < 0.01$ ; \*  $p < 0.05$

Table 11: Régression du montant l'aide totale des parents sur les tranches de revenu parental

ayant un effet similaire sur le montant des aides. À partir de la régression présentée ci-dessus, nous avons défini cinq nouvelles catégories de revenus parentaux :

- < 500€
- 500-2000€
- 2000-3000€
- 3000-5000€
- 5000-8000€
- > 8000€

Nous avons regroupé des niveaux de revenus parfois très éloignés (par exemple 500 € et 2000 €) et donc rassemblé des réalités sociales très différentes. Toutefois, puisque l'essentiel réside dans l'impact de ces niveaux de revenu sur l'aide parentale, c'est la similitude de cet effet qui importe avant tout. Malgré tout, notre sélection demeure arbitraire quant à l'évaluation des similarités observées. L'écart entre l'effet des revenus parentaux inférieurs à 500 € et celui des revenus compris entre 500 € et 1000 € sur l'aide parentale était à peu près le même que celui entre l'effet des revenus compris entre 500 € et 1000 € et celui des revenus compris entre 1500 € et 2000 €. Nous aurions ainsi pu choisir une première tranche de revenus regroupant les parents gagnant moins de 1000 €, plutôt que de constituer une large catégorie englobant ceux dont les revenus se situent entre 500 et 2000 €. Ce choix arbitraire a peut-être légèrement influencé nos résultats, mais cet impact est atténué par le fait que le niveau de revenu parental exerce peu d'influence sur le coefficient associé au sexe dans notre régression.

## IV.4 L'enjeu de prendre une catégorie de référence dans nos régressions

Nous avons systématiquement choisi dans toutes nos régressions comme catégorie de référence pour nos variables catégorielles la catégorie avec le plus d'individus. En effet, le choix des catégories de référence n'affecte pas les prédictions du modèle ou son  $R^2$  mais il permet d'améliorer la précision statistique des coefficients. En effet, plus la catégorie de référence est grosse, plus on compare les autres catégories à une catégorie bien estimée (car plus massive). Les erreurs standards sont donc plus petites et les tests de nullité plus puissants.

## IV.5 Choix de l'hétéroscédasticité et de la non utilisation de clusters

Si on ne fait pas de régression robuste, on suppose que notre échantillon est homoscedastique. C'est-à-dire qu'on suppose que les chocs sur les variables explicatives ont le même effet et la même intensité sur tous les individus. On suppose aussi que les chocs entre individus ne sont pas liés. Cette seconde hypothèse ici tient car, dans l'échantillon étudié, il n'y a aucun jeune avec le même idmen, c'est-à-dire avec le même identifiant de foyer. Ainsi, tous les jeunes ont un foyer indépendant. Il est donc logique de supposer que les chocs sur un jeune n'affectera pas un autre jeune. Cependant l'hypothèse de l'homoscedasticité n'est pas soutenable dans notre cas. C'est pour cela que nous avons recours à des régressions robustes de type HC1 (que nous utilisons car cette correction robuste est utilisée de manière standard, alors que des corrections de type HC2 ou HC3 sont plus appropriées pour des petits échantillons) et pas cluster-robuste.

## IV.6 Motivation du modèle logit et effets marginaux/pourquoi logit plutôt que probit

Le modèle logit est ici utilisé pour tester l'effet des différentes variables sur la marge extensive liée à l'aide que les parents fournissent au jeune. Notre étude se prête parfaitement à l'exercice car nous avons accès à la variable latente (l'aide mensuel) et nous pouvons créer une variable binaire à partir de celle-ci (est-ce que le jeune reçoit un montant strictement positif ou non). C'est d'ailleurs pour cela que nous avons ensuite recours à un modèle tobit qui nous permet d'avoir d'estimer à la fois la marge intensive et la marge extensive. En effet, le modèle logit permet d'estimer la probabilité que le jeune reçoive une aide ou non et la variation de cette probabilité en fonction des différentes caractéristiques du jeune. Nous avons choisi de ne présenter que le modèle logit plutôt que le modèle probit car les résultats étaient quasiment identiques ( $R^2$  quasi égaux, relations d'ordre proportionnelles entre les coefficients, etc...). Pour interpréter le logit, nous avons eu recours aux effets marginaux qui peuvent s'interpréter de manière à la fois cardinale et ordianale alors que les coefficients de la régression ne s'interprète que de manière ordianale. En réalité, les coefficients sont juste les effets marginaux à une transformation linéaire près.

## IV.7 Modèle tobit et effets marginaux

À partir des coefficients du modèle tobit et de leurs écarts-type, il est possible d'estimer des effets marginaux qui nous permettent de comparer dans le cadre du modèle tobit les marges intensive et extensive calculées au lieu de les calculer séparément avec une régression OLS et un modèle logit.

Variable	Coefficient Tobit	Effet marginal moyen	Effet marginal probabilité
(Intercept)	314.296	0.690	0.001
sexe	20.076	0.333	0.000
age	-11.522	13.826	0.022
moins_de_500	-109.427	0.003	0.000
de_2000_3000	60.677	0.175	0.000
de_3000_5000	129.460	0.230	0.000
de_5000_8000	249.797	0.095	0.000
plus_8000	333.781	0.022	0.000
sit_emploi	-284.525	0.093	0.000
sit_apprenti	-173.649	0.045	0.000
sit_chômage	-118.771	0.070	0.000
sit_autre	-147.199	0.020	0.000
semi_cohab	265.607	0.314	0.000
non_cohab	76.519	0.106	0.000

L'effet marginal en termes de probabilité s'interprète de la même manière que l'effet marginal logit. Il représente la marge extensive. L'effet marginal moyen mesure l'impact moyen de la variable explicative sur le montant réel de l'aide parentale, en tenant compte des valeurs latentes négatives estimées par le modèle Tobit, tout en considérant que ces valeurs négatives sont observées comme zéro. Toutefois, les valeurs obtenues sont très étranges, ce qui nous a amenés à penser que notre calcul n'était pas correct, sans toutefois avoir eu le temps d'investiguer suffisamment en profondeur le problème sous-jacent. D'une part, les effets marginaux moyens sont tous compris entre 0 et 1, alors qu'ils devraient s'exprimer en montants en euros. Or, d'après le calcul des autres coefficients, on sait que les effets en marge intensive sont importants et devraient pousser cet effet marginal moyen vers le haut. D'autre part, les coefficients marginaux exprimés en termes de probabilités sont également très faibles comparés aux coefficients logit observés, ce qui nous semble incohérent.

## IV.8 La décomposition de Gelbach

Covariable	coefficient OLS (1)	coefficient OLS (2)	coefficient Gelbach
age	-2.883	0.016	-0.045
de_2000_3000	33.993	-0.013	-0.426
de_3000_5000	91.563	-0.009	-0.819
de_5000_8000	196.617	0.001	0.122
moins_de_500	-57.846	0.003	-0.156
non_cohab	98.190	0.085	8.346
plus_8000	278.152	-0.011	-2.988
semi_cohab	251.705	0.013	3.147
sit_apprenti	-139.019	-0.015	2.059
sit_autre	-118.516	0.013	-1.577
sit_chômage	-97.496	-0.036	3.493
sit_emploi	-188.217	-0.044	8.328

La décomposition de Gelbach est une méthode qui permet de quantifier la contribution individuelle de chaque variable de contrôle à la variation du coefficient d'intérêt. En effet, il n'est pas pertinent d'ajouter ces variables une par une dans la régression, car cela introduirait un biais lié à l'ordre d'inclusion des variables.

Principe de la méthode:

1) **déterminer  $\beta_S$ , coefficient associé au sexe dans la régression courte et les coefficients de la variable d'intérêt mais aussi de toutes les variables de contrôle.** ( donne dans le tableau ci-dessus la colonne "coefficient OLS (1)")

2) **Régresser chaque variable de contrôle sur le sexe.**( donne dans le tableau ci-dessus la colonne "coefficient OLS (2)")

On obtient ainsi un coefficient  $\Gamma_C$  tel que, soit notons  $C$  la variable de contrôle, on a:

$$C = \Gamma_C * sexe + \varepsilon \quad (1)$$

3) **En remplaçant dans la régressions avec la variable d'intérêt et toutes les variables de contrôle, on obtient enfin la contribution relative à l'évolution du coefficient associé au sexe de chaque variable.**



Pour plus de simplicité, on suppose ici qu'il y a trois variables de contrôle notées C\_1, C\_2 et C\_3. En remplaçant chaque variable par leur régression respective sur la variable sexe, on obtient:

$$\text{aide} = \beta_L \cdot \text{sexe} + \beta_{C1} \cdot \Gamma_{C1} \cdot \text{sexe} + \beta_{C2} \cdot \Gamma_{C2} \cdot \text{sexe} + \beta_{C3} \cdot \Gamma_{C3} \cdot \text{sexe} + \beta_{C1}\varepsilon_1 + \beta_{C2}\varepsilon_2 + \beta_{C3}\varepsilon_3 + \eta$$

d'où :

$$\text{aide} - \text{aide} = \beta_L - \beta_S + \beta_{C1} \cdot \Gamma_{C1} + \beta_{C2} \cdot \Gamma_{C2} + \beta_{C3} \cdot \Gamma_{C3} + \nu$$

avec  $\nu = \beta_{C1}\varepsilon_1 + \beta_{C2}\varepsilon_2 + \beta_{C3}\varepsilon_3 + \eta$

Par conséquent, in fine on obtient:

$$\beta_S - \beta_L = \beta_{C1} * \Gamma_{C1} + \beta_{C2} * \Gamma_{C2} + \beta_{C3} * \Gamma_{C3} + \nu \quad (2)$$

**La contribution marginale de chaque variable de contrôle est donc mesurée par  $\beta_C * \Gamma_C$ .** La différence entre le coefficient associé au sexe dans le modèle simple et celui associé à cette même variable d'intérêt dans le modèle plus long est égale à la somme des contributions des variables de contrôle et d'un résidu exogène, non expliqué.

## IV.9 Le choix des types d'aide à étudier

Pour sélectionner les différents types d'aides à étudier, nous avons lu attentivement le dictionnaire des variables et nous avons ensuite regroupé les différentes aides dans les huit catégories suivantes : les aides au logement, les aides aux loisirs, les aides concernant la santé du jeune, les aides pour la scolarité du jeune, les aides pour la communication, les aides pour les transports, les aides pour les vêtements et l'habillement et enfin, les aides pour les voyages. Ces catégories nous ont permis d'étudier un large éventail d'aides tout en conservant des échantillons suffisamment larges (dans une grande partie des cas) pour proposer des résultats intéressants.

Aussi les différentes variables d'aides binaires ont été construites de la façon suivante :

- L'aide binaire pour le logement a été construite grâce aux variables catégorielles `aideloy1_c`, `aideloy2_c`, `aidlogp3`, `aidlogp2`, `aidechgepar1`, `aidechgepar2`, `achlog1pm` et `achlog2pm`. Suivant cela, il a été considéré que le jeune recevait une aide au logement dans les cas suivants : si les parents aident le jeune pour le loyer, les charges, le chauffage, l'électricité, l'assurance etc. du logement 1/1 bis ou du logement 2 ; si les parents ont aidé le jeune à aménager son logement indépendant ; si les parents aident ou ont aidé le jeune à financer le loyer de son logement ; si les parents aident le jeune pour les charges, le chauffage, l'électricité etc. du logement 1/1 bis ou du logement 2 ; et enfin, si les parents l'ont aidé pour l'achat du logement 1/1 bis ou du logement 2.
- L'aide binaire pour les loisirs a été construite grâce aux variables binaires `aideloisirs_a`, `aideloisirs_b`, `aideloisirs_c` et `aideloisirs_d`. Suivant cela, il a été considéré que le jeune recevait une aide pour sa consommation de loisirs dans les cas suivants : si les parents l'aident à payer des licences de club sportif, un abonnement à un club de gym ou de fitness ; si les parents l'aident à payer des sorties au cinéma, au théâtre, aux concerts, au musée, dans des bars ou en boîte de nuit ; si les parents l'aident

à s'acheter des livres, de la musique, des vidéos, des DVD ou des jeux vidéo ; ou autre.

- L'aide binaire pour la santé a été construite grâce à la variable catégorielle `aidemc_r` et suivant cela il a été considéré que le jeune recevait une aide pour ses dépenses liées à la santé si ses parents l'aident à payer sa complémentaire santé.
- L'aide binaire pour la scolarité a été construite à partir de la variable catégorielle `aidscol_r` et suivant cela, il a été considéré que le jeune recevait une aide pour les dépenses liées à sa scolarité si ses parents prennent en charge tout ou une partie des frais de scolarité ou universitaires.
- L'aide binaire pour la communication a été construite à partir de la variable catégorielle `aidcomm` et suivant cela, il a été considéré que le jeune recevait une aide pour ses dépenses de communication si ses parents l'aident pour les dépenses de téléphone, d'abonnement téléphone et/ou d'internet.
- L'aide binaire pour les transports a été construite à partir des variables `aideabonne`, `aideessence`, `aquivoiture`, `aideassur` et `billets_c`. Suivant cela, il a été considéré que le jeune recevait une aide pour les transports dans les cas suivants : si ses parents payent directement en totalité les abonnements pour les transports en commun ou s'ils donnent au jeune de l'argent pour payer ses abonnements en plus de leur éventuelle aide financière régulière ; si ses parents l'aident pour les dépenses d'essence et d'entretien de son véhicule ; si ses parents l'ont aidé à acheter le véhicule qu'il utilise le plus souvent ou s'ils le lui ont donné ou prêté ; si ses parents lui donnent de l'argent pour payer l'assurance de son véhicule ou si celle-ci est au nom de l'un de ses deux parents ; et enfin, si ses parents lui payent régulièrement des billets de train (hors abonnement).
- L'aide binaire pour les vêtements et l'habillement a été construite à partir de la variable catégorielle `aidevet` et suivant cela, il a été considéré que le jeune recevait une aide pour ses dépenses en vêtements si ses parents l'aident à acheter des vêtements pour lui ou ses enfants souvent ou parfois.
- L'aide binaire pour les voyages a été construite à partir de la variable `aidvoyage` et suivant cela, il a été considéré que le jeune recevait une aide pour voyager si ses parents l'aident pour financer des voyages et séjours.

De la même façon, les différentes variables de montant ont été construites de la façon suivante :

- La variable du montant de l'aide perçue pour le logement a été construite à partir des variables quantitatives `mtchgepar1`, `mtchgepar2`, `mensualitepm1`, `mensualitepm2`, `ctrlog1par_r` et `ctrlog2par_r`. Suivant cela, il a été considéré que le montant reçu par le jeune de la part de ses parents pour son logement était la somme : du montant mensuel payé par ses parents pour le logement 1/1 bis et du montant mensuel payé par ses parents pour le logement 2.
- La variable du montant de l'aide perçue pour les loisirs a été construite à partir de la variable `mtaideloisirs_r`. Suivant cela, il a été considéré que le montant reçu par le jeune de la part de ses parents pour ses dépenses de loisirs était le montant mensuel des dépenses de loisirs payées par ses parents.

- La variable du montant de l'aide perçue pour la santé a été construite à partir de la variable `mtaideamcee_c`. Suivant cela, il a été considéré que le montant reçu par le jeune de la part de ses parents pour ses dépenses de santé était le montant de l'aide des parents pour la complémentaire santé. A noter que nous avons considéré que les jeunes qui ne savaient pas répondre à la question, recevait un montant nul pour cette aide.
- La variable du montant de l'aide perçue pour la scolarité a été construite à partir de la variable `mtaidscol_r`. Suivant cela, il a été considéré que le montant reçu par le jeune de la part de ses parents pour ses dépenses liées à la scolarité était le montant annuel payé par ses parents pour les frais de scolarité.
- La variable du montant de l'aide perçue pour la communication a été construite à partir de la variable `mtaidcomm_r`. Suivant cela, il a été considéré que le montant reçu par le jeune de la part de ses parents pour ses dépenses de communication était le montant mensuel payé par ses parents pour ses dépenses de téléphone, d'abonnement téléphone et/ou d'internet.
- La variable du montant de l'aide perçue pour les transports a été construite à partir des variables quantitatives `mtbillets_r`, `mtaidessence_r`, `mtaidassur_r` et `mtaidabonne_r`. Suivant cela, il a été considéré que le montant reçu par le jeune de la part de ses parents pour ses déplacements était la somme : du montant mensuel de l'achat de billets de train par ses parents ; du montant mensuel moyen de l'aide versée par ses parents pour l'essence et l'entretien de son véhicule ; du montant annuel de l'assurance prise en charge par ses parents (divisé par douze) ; et enfin du montant mensuel moyen payé par ses parents pour les abonnements aux transports en commun.
- La variable du montant de l'aide perçue pour les vêtements et l'habillement a été construite à partir de la variable quantitative `mtaidevet` et suivant cela, il a été considéré que le montant reçu par le jeune de la part de ses parents pour ses dépenses en vêtements était le montant annuel dépensé par ses parents pour les achats de vêtements.
- La variable du montant de l'aide perçue pour les voyages a été construite à partir de la variable quantitative `mtvoyage_r` et suivant cela il a été considéré que le montant reçu par le jeune de la part de ses parents pour ses dépenses en voyages était le montant annuel de l'aide de ses parents pour les voyages et séjours.

Et pour les variables de contrôle, nous avons procédé de la façon suivante :

- Nous avons créé la variable `cohab.binaire` à partir de la variable `cohab`, et nous avons considéré qu'un jeune vivait chez ses parents lorsqu'il déclarait résider systématiquement avec au moins un de ses parents.
- Nous avons créé les variables binaires `sitprinc_Emploi`, `sitprinc_Apprentissage`, `sitprinc_Chomage` et `sitprinc_Autre`, à partir de la variable `sitprinc`. Et nous avons pris la situation en étude comme référence car la majorité des jeunes font partie de cette catégorie.
- La variable `aide_partenaire_logement` a été créée à partir de la variable `apportc_j_a`, et nous avons considéré que les jeunes n'ayant pas répondu à la question ne recevait pas d'aide de la part de leur partenaire éventuel.

## IV.10 Suite des régressions logistiques par type d'aide

### Aide binaire pour les dépenses de santé

Bien qu'ils ne soient pas significatifs, les résultats des régressions logistiques effectuées sur l'aide à la marge extensive pour les dépenses de santé sont intéressants à étudier. Il semblent montrer que toutes choses égales par ailleurs, le fait d'être une femme augmente la probabilité de recevoir cette aide à hauteur de 10,5 points de pourcentage. Cela n'est pas négligeable, toutefois, le manque de significativité du coefficient estimé dans la régression ne nous permet pas de conclure sur ce point.

Table 12: Effets marginaux loisirs - Modèle avec l'état de santé

Variable	Effet marginal	p-value
agej_scaled	-0.0042	0.4978
cohab_binaire	-0.0269	0.02319
etatsante_AsezBon	-0.0065	0.7332
etatsante_Bon	-0.0305*	0.01428
etatsante_Mauvais	-0.0713	0.1706
femme	0.0109	0.3418
revenu_r_scaled	0.026***	1.846e-06
sitprinc_Apprentissage	-0.1875***	3.566e-09
sitprinc_Autre	-0.1325***	0.0001544
sitprinc_Chomage	-0.1123***	1.122e-07
sitprinc_Emploi	-0.172***	1.141e-17

Note: \*p<0.05; \*\*p<0.01; \*\*\*p<0.001

### Aide binaire pour les dépenses liées à la scolarité

De la même façon, les résultats des régressions logistiques effectuées sur l'aide à la marge extensive pour les dépenses liées à la scolarité ne sont pas significatifs (la p-value de l'effet marginal moyen pour la variable "femme" est supérieure à 0,05). Nous remarquons de plus que l'effet estimé est très faible : toutes choses égales par ailleurs, le fait d'être une femme augmenterait en moyenne, la probabilité de recevoir une aide pour la scolarité de 1,8 point de pourcentage. Nous pouvons donc supposer que le fait d'être une femme n'a pas nécessairement d'effet sur le fait de recevoir ou non une aide de la part de ses parents pour financer ses frais de scolarité.

Table 13: Effets marginaux scolarité - Modèle avec les variables de contrôle usuelles

Variable	Effet marginal	p-value
agej_scaled	-0.036***	1.21e-06
cohab_binaire	0.013	0.303
femme	0.018	0.161
revenu_r_scaled	0.116***	4.78e-58
sitprinc_Apprentissage	-0.151***	2.03e-12
sitprinc_Autre	-0.288***	2.59e-12
sitprinc_Chomage	-0.662***	1.5e-24
sitprinc_Emploi	-0.53***	1.88e-55

Note: \*p<0.05; \*\*p<0.01; \*\*\*p<0.001

### Aide binaire pour les dépenses de communication (téléphonie et internet)

En ce qui concerne l'aide des parents pour les dépenses de communication, comme pour l'aide pour les dépenses de scolarité, l'effet marginal moyen estimé est à la fois faible et peu représentatif, ce qui nous pousse à prédire que le sexe n'a pas d'effet à la marge extensive sur l'aide parentale pour la communication. Et s'il y a un effet du sexe, il est en faveur des femmes.

Table 14: Effets marginaux communication - Modèle avec les variables de contrôle usuelles

Variable	Effet marginal	P_value
agej_scaled	-0.132***	2.5e-76
cohab_binaire	0.157***	1.66e-25
femme	0.017	0.255
revenu_r_scaled	0.066***	5.44e-10
sitprinc_Apprentissage	-0.216***	4.32e-12
sitprinc_Autre	-0.128***	0.000299
sitprinc_Chomage	-0.095***	1.65e-05
sitprinc_Emploi	-0.286***	3.68e-62

Note: \*p<0.05; \*\*p<0.01; \*\*\*p<0.001

### Aide binaire pour les dépenses de transports

Pour ce qui est de l'aide pour les transports, les femmes semblent être favorisées. En effet, avec une p-value inférieure à 0.05, l'effet marginal moyen du sexe implique que : toutes choses égales par ailleurs, le fait d'être une femme augmente en moyenne de 3,6 points de pourcentage la probabilité de recevoir une aide pour les transports de la part de ses parents.

Table 15: Effets marginaux transports - Modèle sans l'aide du partenaire

Variable	Effet marginal	p-value
agej_scaled	-0.018*	0.0474
cohab_binaire	0.155***	7.96e-21
femme	0.036*	0.024
revenu_r_scaled	0.106***	1.26e-18
sitprinc_Apprentissage	-0.141***	2.41e-05
sitprinc_Autre	-0.238***	3.34e-09
sitprinc_Chomage	-0.2***	4.42e-16
sitprinc_Emploi	-0.171***	1.38e-15
train_c	-0.09*	0.0353

Note: \*p<0.05; \*\*p<0.01; \*\*\*p<0.001

### Aide binaire pour les dépenses liées aux vêtements

En revanche, concernant l'aide parentale pour l'achat de vêtements, le coefficient marginal moyen est très faible (associé à une augmentation moyenne de la probabilité de moins d'un point de pourcentage). Et cela, avec une p-value supérieure à 0,9. Nous ne pouvons donc pas conclure quant à l'effet d'être une femme sur le fait de recevoir ou non une aide pour s'acheter des vêtements de la part de ses parents.

Table 16: Effets marginaux vêtements - Modèle avec les variables de contrôle usuelles

Variable	Effet marginal	p-value
agej_scaled	-0.093***	8.66e-29
cohab_binaire	0.156***	3.15e-22
femme	-0.002	0.922
revenu_r_scaled	0.056***	3.82e-09
sitprinc_Apprentissage	-0.153***	2.91e-06
sitprinc_Autre	-0.058	0.156
sitprinc_Chomage	-0.064*	0.0106
sitprinc_Emploi	-0.227***	9.96e-30

Note: \*p<0.05; \*\*p<0.01; \*\*\*p<0.001

### Aide binaire pour les dépenses liées aux voyages et séjours

“ Comme nous l'avons remarqué pour l'aide concernant la scolarité et l'aide concernant la communication, l'effet marginal moyen estimé du fait d'être une femme sur le fait de recevoir ou non une aide pour voyager de la part de ses parents n'est pas significatif. Toutefois, il est positif, ce qui laisse supposer qu'une étude menée sur un plus grand nombre d'individus pourrait nous permettre de conclure que les femmes sont plus souvent aidées par leurs parents pour voyager que les hommes.

Table 17: Effets marginaux voyages - Modèle avec les variables de contrôle usuelles

Variable	Effet marginal	p-value
agej_scaled	-0.015	0.0516
cohab_binaire	0.026	0.0526
femme	0.021	0.116
revenu_r_scaled	0.06***	8.43e-20
sitprinc_Apprentissage	-0.097***	0.000316
sitprinc_Autre	-0.095**	0.00951
sitprinc_Chomage	-0.132***	4.1e-08
sitprinc_Emploi	-0.207***	6.02e-22

Note: \*p<0.05; \*\*p<0.01; \*\*\*p<0.001

## IV.11 Informations complémentaires sur l'étude menée sur les écarts dans les montants reçus par les jeunes

### Attrition pour chez les jeunes déclarant une aide au logement

	Aide = 0	Aide = 1
Montant = 0	3072	1441
Montant ≠ 0	0	1263

On remarque tout d'abord qu'aucun jeune ayant déclaré ne pas recevoir d'aide pour le logement de la part de ses parents n'a indiqué de montant, ce qui était attendu. Toutefois, nous remarquons également que parmi les 3 072 jeunes ayant déclaré recevoir une aide, seuls 1 263 ont renseigné un montant.

Pour évaluer l'importance de ce phénomène d'attrition dans l'étude que nous avons menée, nous avons décidé de construire un tableau montrant la distribution dans les deux groupes. Ce tableau nous montre que l'attrition est sélective : par exemple, toutes les personnes ayant renseigné un montant vivent dans un logement autre que celui de leurs parents (au moins une partie du temps), ce qui est loin d'être le cas pour le groupe des jeunes ayant déclaré une aide mais pas de montant.

	Montant = 0	Montant ≠ 0
Femme	48.36	56.85
Âge Moyen	21.09	20.15
Revenu Moyen Parents	2978.39	4042.88
Emploi	33.47	2.88
Apprentissage	6.79	5.69
Chomage	17.97	1.13
Autre	5.21	1.27
Cohabitation	65.22	0
APL	48.22	63.25
Aide Partenaire	38.5	23.3

Cette attrition peut venir du fait que certaines personnes ont déclaré recevoir une aide alors que ça n'était pas le cas et n'ont donc pas renseigné de montant. Dans ce cas, l'attrition ne pose aucun problème puisque nous avons exclu ces individus de notre régression (les considérant ainsi comme les jeunes n'ayant pas reçu l'aide). Toutefois, dans le cas où ces individus ont réellement reçu une aide, mais n'ont simplement pas renseigné de montant alors l'attrition sélective peut devenir une grande faiblesse de l'analyse. Ainsi, nous devons bien garder à l'esprit que les résultats concernant l'aide à la marge

intensive pour le logement, les loisirs et les transports peuvent être en partie corrompus par ce phénomène d'attrition sélective.

### Régressions n'ayant pas pu être effectuées, faute d'effectifs

Comme nous l'avons évoqué précédemment, nous avons décidé de ne pas effectuer les régressions si moins de 500 jeunes présents dans la base ont déclaré un montant pour la variable d'aide en question. Pour cette raison, nous n'avons pas réalisé les régressions linéaires sur les variables d'aide en montant pour la santé et la communication pour lesquelles nous avons respectivement 477 et 128 observations positives.

### Aide en montant pour les loisirs

Les régressions linéaires effectuées pour l'aide concernant les loisirs à la marge intensive nous donnent des coefficients non significatifs (p-value supérieure à 0,05). De plus, les coefficients sont plutôt faibles. Notons toutefois que les coefficients sont positifs, ce qui signifie que : les jeunes hommes recevrait plus souvent une aide pour leurs dépenses de loisirs de la part de leurs parents mais celle-ci serait moins élevée que l'aide perçue par les jeunes femmes.

Table 18: Régressions linéaires pour l'aide aux loisirs

	<i>Dependent variable:</i>	
	montant_aide_loisirs	
	Sans contrôle	Contrôles usuels
	(1)	(2)
femme	1.556	2.360
revenu_r		0.004*
agej		1.449
sitprinc_Emploi		6.094
sitprinc_Apprentissage		37.499
sitprinc_Chomage		-0.334
sitprinc_Autre		-4.613
cohab_binaire		3.268
Constant	54.763***	5.984
Observations	1,311	1,311
R <sup>2</sup>	0.0001	0.021
Adjusted R <sup>2</sup>	-0.001	0.015
Residual Std. Error	2,551.237 (df = 1309)	2,530.997 (df = 1302)
F Statistic	0.110 (df = 1; 1309)	3.516*** (df = 8; 1302)

Note: \*p<0.05; \*\*p<0.01; \*\*\*p<0.001



## Aide en montant pour la scolarité

	Aide = 0	Aide = 1
Montant = 0	4146	0
Montant ≠ 0	0	1630

Contrairement aux aides en montant pour le logement et les loisirs, nous remarquons que l'aide pour la scolarité ne subit pas d'attrition.

Encore une fois, les résultats des deux régressions manquent de significativité. Comme nous pouvons l'observer dans la table de régression qui suit, les coefficients estimés associés à la variable femme ont tous les deux une p-value supérieure à 0,05. Toutefois, nous pouvons noter que les deux coefficients sont négatifs et très éloignés de zéro. Cela tend à montrer que les jeunes femmes qui reçoivent une aide pour la scolarité de la part de leurs parents sont désavantagées (à hauteur de 120€ par an toutes choses égales par ailleurs) par rapport aux jeunes hommes dans la même situation. Cependant, la non significativité de nos résultats ne nous permet pas de conclure sur ce point.

Table 19: Régressions linéaires pour l'aide concernant la scolarité

	<i>Dependent variable:</i>	
	montant_aide_scolarité	
	Sans contrôle	Contrôles usuels
	(1)	(2)
femme	-192.449	-119.549
revenu_r		0.186***
agej		206.409***
sitprinc_Emploi		-1.843
sitprinc_Apprentissage		-60.864
sitprinc_Chomage		-940.867***
sitprinc_Autre		-594.591
cohab_binaire		-19.336
Constant	1,564.025***	-3,384.534***
Observations	1,541	1,541
R <sup>2</sup>	0.002	0.075
Adjusted R <sup>2</sup>	0.001	0.070
Residual Std. Error	71,169.270 (df = 1539)	68,675.370 (df = 1532)
F Statistic	2.457 (df = 1; 1539)	15.430*** (df = 8; 1532)
<i>Note:</i> *p<0.05; **p<0.01; ***p<0.001		

## Aide en montant pour les transports

Table 20: Régressions linéaires pour l'aide concernant les transports

	<i>Dependent variable:</i>	
	montant_aide_transports	
	Sans contrôle	Contrôles usuels
	(1)	(2)
femme	5.968	5.830
revenu_r		-0.0001
agej		1.532
sitprinc_Emploi		-6.419
sitprinc_Apprentissage		-7.671
sitprinc_Chomage		-2.311
sitprinc_Autre		12.523
cohab_binaire		0.706
Constant	65.790***	35.411
Observations	1,404	1,404
R <sup>2</sup>	0.003	0.009
Adjusted R <sup>2</sup>	0.002	0.003
Residual Std. Error	1,697.568 (df = 1402)	1,696.742 (df = 1395)
F Statistic	3.861** (df = 1; 1402)	1.529 (df = 8; 1395)
<i>Note:</i> *p<0.05; **p<0.01; ***p<0.001		

Ici, comme pour les régressions précédentes, les résultats ne sont pas significatifs. Les deux coefficients sont toutefois positifs et proches de 5. Ce qui laisse sous-entendre un effet du sexe sur le montant reçu de 5€ par mois supplémentaires pour les femmes toutes choses égales par ailleurs.

## Aide en montant pour les vêtements

En ce qui concerne l'aide pour les vêtements, comme pour l'aide concernant la scolarité, il n'y pas de problème d'attrition. Cependant, encore une fois, le manque de significativité des résultats ne nous permet pas de conclure efficacement sur l'effet du sexe sur le montant reçu pour cette aide.

Table 21: Régressions linéaires pour l'aide concernant les vêtements

	<i>Dependent variable:</i>	
	montant_aide_vêtements	
	Sans contrôle	Contrôles usuels
	(1)	(2)
femme	-1.071	4.966
revenu_r		0.029***
agej		-17.759***
sitprinc_Emploi		-39.850
sitprinc_Apprentissage		-20.076
sitprinc_Chomage		-48.531*
sitprinc_Autre		-26.139
cohab_binaire		24.876*
Constant	288.058***	549.524***
Observations	3,007	3,007
R <sup>2</sup>	0.00000	0.057
Adjusted R <sup>2</sup>	-0.0003	0.054
Residual Std. Error	10,630.940 (df = 3005)	10,337.600 (df = 2998)
F Statistic	0.007 (df = 1; 3005)	22.496*** (df = 8; 2998)
<i>Note:</i> *p<0.05; **p<0.01; ***p<0.001		