

Valentin Barthel

15 avril 2024

Table des matières

1 Question 1 : Reprendre les données des exercices 1 et 2 et ajouter la variable lrd0, correspondant au logarithme du revenu d'inactivité (revenu perçu si la personne ne travaille pas). Créer la variable "femme"=Isex=2 ainsi que la variable "homme"=Isex=1 (déjà fait dans l'exercice 1, pour ceux qui l'ont traité)	2
2 Question 2 : Régresser (indépendamment) par MCO l'activité principale (actifp) et l'activité secondaire (actifs) sur les variables sexe et lrd0. Interpréter et commenter les résultats. Idem avec femme et lrd0, puis avec homme et lrd0. Commenter les similitudes et différences.	3
3 Question 3 : Reprendre la question avec un modèle logit puis avec un modèle Probit. Commenter les similitudes et différences par rapport aux MCO.	12
4 Question 4 : (papier-crayon)	20
5 Question 5 : Calculer ces différentes probabilités pour chaque ligne de l'échantillon et tracer sur un même graphique toutes les courbes de probabilité d'activité principale en fonction du log-revenu d'inactivité. Idem pour la probabilité d'activité secondaire. Vérifier la cohérence avec la question précédente.	25
6 Question 6 :	30

1 Question 1 : Reprendre les données des exercices 1 et 2 et ajouter la variable lrd0, correspondant au logarithme du revenu d'inactivité (revenu perçu si la personne ne travaille pas). Créer la variable "femme"=Isexec=2 ainsi que la variable "homme"=Isexec=1 (déjà fait dans l'exercice 1, pour ceux qui l'ont traité)

```
----- QUESTION 1 -----  
/* Définition de la librairie et appel de la base  
"merged_panelEX2" enregistrée dans l'exercice 2*/  
libname maLib '/home/u63824485/sasuser.v94';  
  
data merged_panelex2;  
  set maLib.merged_panelEX2;  
run;  
  
  /* Création de la variable lrd0, la mise en logarithme de la variable rd0*/  
data merged_panelex3;  
  set merged_panelex2;  
  if rd0 > 0 then lrd0 = log(rd0);  
  else lrd0 = .;  
run;  
  
/*L'exercice 1 ayant été traité, la base contient la variable 'sexen' pour laquelle elle est prend comme valeur 1 si l'individu observé est un homme et 2 si l'individu observé est une femme.*/
```

2 Question 2 : Régresser (indépendamment) par MCO l'activité principale (actifp) et l'activité secondaire (actifs) sur les variables sexe et lrd0. Interpréter et commenter les résultats. Idem avec femme et lrd0, puis avec homme et lrd0. Commenter les similitudes et différences.

```
/*-----QUESTION 2-----*/
/*Régression actifp et actifs expliqué par sexeN */
proc reg data=merged_panelex3;
model actifp = sexeN lrd0;
title "Régression de actifp sur sexeN et lrd0";
run;
quit;

proc reg data=merged_panelex3;
model actifs = sexeN lrd0;
title "Régression de actifs sur sexeN et lrd0";
run;
quit;

/*Régression actifp et actifs expliqué par sexeHomme */
proc reg data=merged_panelex3;
model actifp = sexeHomme lrd0;
title "Régression de actifp sur sexeHomme et lrd0";
run;
quit;

proc reg data=merged_panelex3;
model actifs = sexeHomme lrd0;
title "Régression de actifs sur sexeHomme et lrd0";
run;
quit;

/*Régression actifp et actifs expliqué par sexeFemme */
proc reg data=merged_panelex3;
model actifp = sexeFemme lrd0;
title "Régression de actifp sur sexeFemme et lrd0";
run;
```

```
quit;

proc reg data=merged_panelex3;
  model actifs = sexeFemme lrd0;
  title "Régression de actifs sur sexeFemme et lrd0";
run;
quit;
```

Régression de actifp sur sexeN et lrd0

La procédure REG
Modèle : MODEL1
Variable dépendante : actifp

Nb d'observations lues	10548
Nb d'obs. utilisées	10548

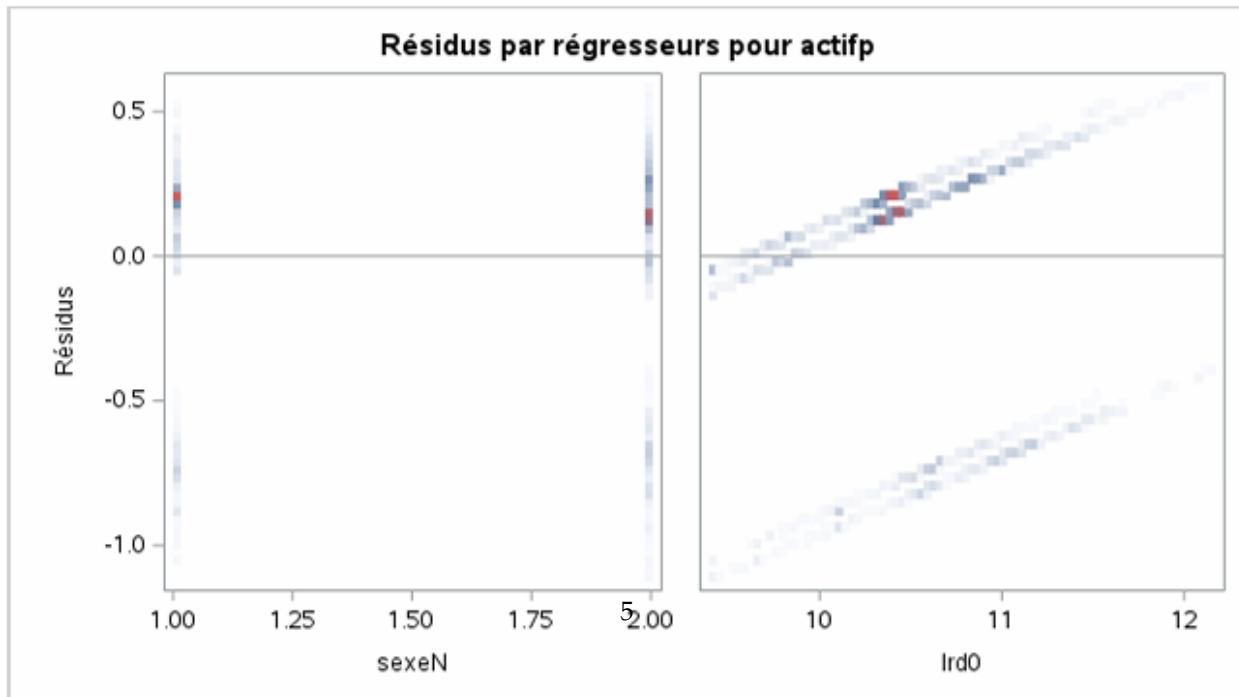
Analyse de variance					
Source	DDL	Somme des carrés	Moyenne quadratique	Valeur F	Pr > F
Modèle	2	154.54008	77.27004	528.90	<.0001
Erreur	10545	1540.56534	0.14609		
Total sommes corrigées	10547	1695.10542			

Root MSE	0.38222	R carré	0.0912
Moyenne dépendante	0.79882	R car. ajust.	0.0910
Coeff Var	47.84818		

Paramètres estimés					
Variable	DDL	Valeur estimée des paramètres	Erreur type	Valeur du test t	Pr > t
Intercept	1	3.48277	0.08340	41.76	<.0001
sexeN	1	0.06835	0.00793	8.62	<.0001
lrd0	1	-0.26606	0.00818	-32.52	<.0001

Régression de actifp sur sexeN et lrd0

La procédure REG
Modèle : MODEL1
Variable dépendante : actifp



Régression de actifs sur sexeN et lrd0

La procédure REG
Modèle : MODEL1
Variable dépendante : actifs

Nb d'observations lues	10548
Nb d'obs. utilisées	10548

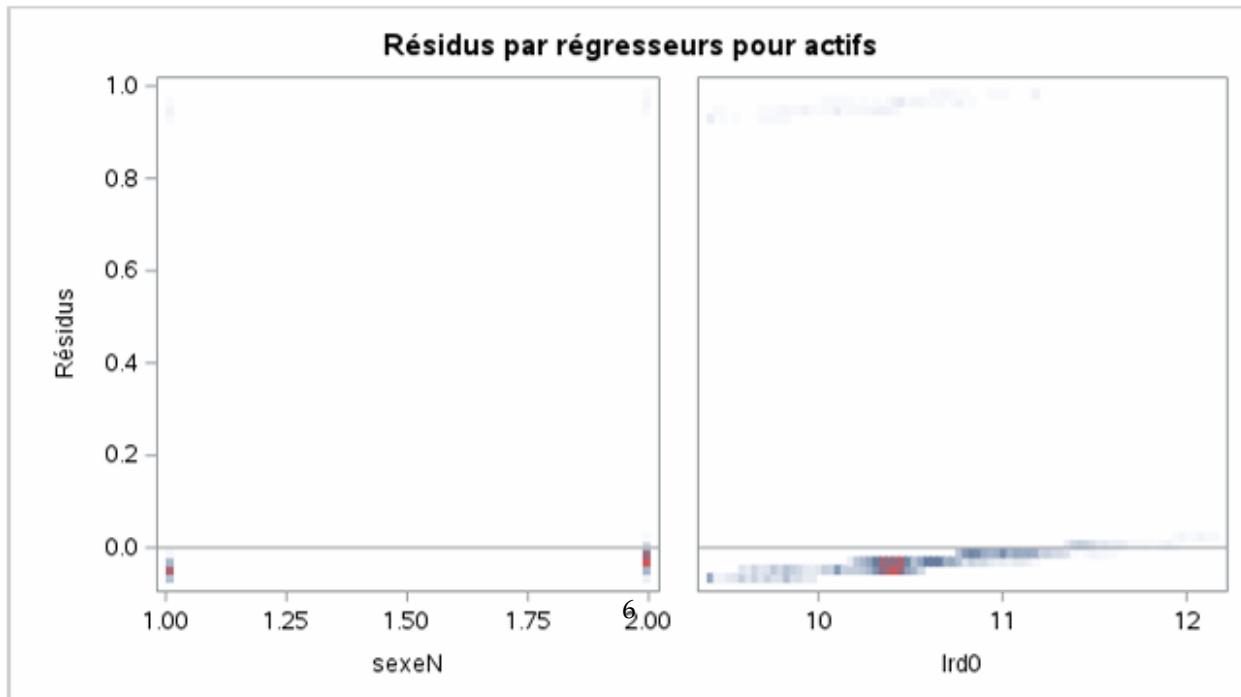
Analyse de variance					
Source	DDL	Somme des carrés	Moyenne quadratique	Valeur F	Pr > F
Modèle	2	2.97787	1.48894	49.14	<.0001
Erreur	10545	319.50933	0.03030		
Total sommes corrigées	10547	322.48720			

Root MSE	0.17407	R carré	0.0092
Moyenne dépendante	0.03157	R car. ajust.	0.0090
Coeff Var	551.37157		

Paramètres estimés					
Variable	DDL	Valeur estimée des paramètres	Erreur type	Valeur du test t	Pr > t
Intercept	1	0.36649	0.03798	9.65	<.0001
sexeN	1	-0.01201	0.00361	-3.33	0.0009
lrd0	1	-0.03006	0.00373	-8.07	<.0001

Régression de actifs sur sexeN et lrd0

La procédure REG
Modèle : MODEL1
Variable dépendante : actifs



Régression de actifp sur sexeHomme et lrd0

La procédure REG
Modèle : MODEL1
Variable dépendante : actifp

Nb d'observations lues	10548
Nb d'obs. utilisées	10548

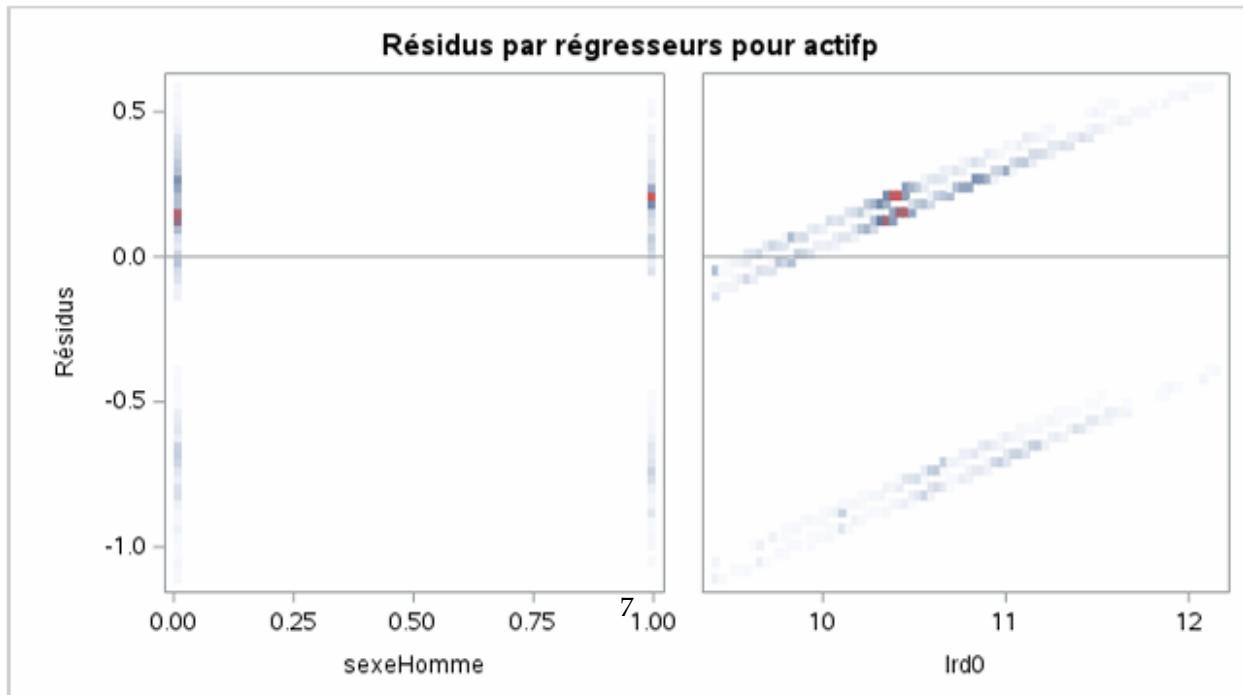
Analyse de variance					
Source	DDL	Somme des carrés	Moyenne quadratique	Valeur F	Pr > F
Modèle	2	154.54008	77.27004	528.90	<.0001
Erreur	10545	1540.56534	0.14609		
Total sommes corrigées	10547	1695.10542			

Root MSE	0.38222	R carré	0.0912
Moyenne dépendante	0.79882	R car. ajust.	0.0910
Coeff Var	47.84818		

Paramètres estimés					
Variable	DDL	Valeur estimée des paramètres	Erreur type	Valeur du test t	Pr > t
Intercept	1	3.61947	0.08689	41.66	<.0001
sexHomme	1	-0.06835	0.00793	-8.62	<.0001
lrd0	1	-0.26606	0.00818	-32.52	<.0001

Régression de actifp sur sexeHomme et lrd0

La procédure REG
Modèle : MODEL1
Variable dépendante : actifp



Régression de actifs sur sexeHomme et lrd0

La procédure REG
Modèle : MODEL1
Variable dépendante : actifs

Nb d'observations lues	10548
Nb d'obs. utilisées	10548

Analyse de variance					
Source	DDL	Somme des carrés	Moyenne quadratique	Valeur F	Pr > F
Modèle	2	2.97787	1.48894	49.14	<.0001
Erreur	10545	319.50933	0.03030		
Total sommes corrigées	10547	322.48720			

Root MSE	0.17407	R carré	0.0092
Moyenne dépendante	0.03157	R car. ajust.	0.0090
Coeff Var	551.37157		

Paramètres estimés					
Variable	DDL	Valeur estimée des paramètres	Erreur type	Valeur du test t	Pr > t
Intercept	1	0.34246	0.03957	8.65	<.0001
sexHomme	1	0.01201	0.00361	3.33	0.0009
lrd0	1	-0.03006	0.00373	-8.07	<.0001

Régression de actifs sur sexeHomme et lrd0

La procédure REG
Modèle : MODEL1
Variable dépendante : actifs



Régression de actifp sur sexeFemme et lrd0

La procédure REG
Modèle : MODEL1
Variable dépendante : actifp

Nb d'observations lues	10548
Nb d'obs. utilisées	10548

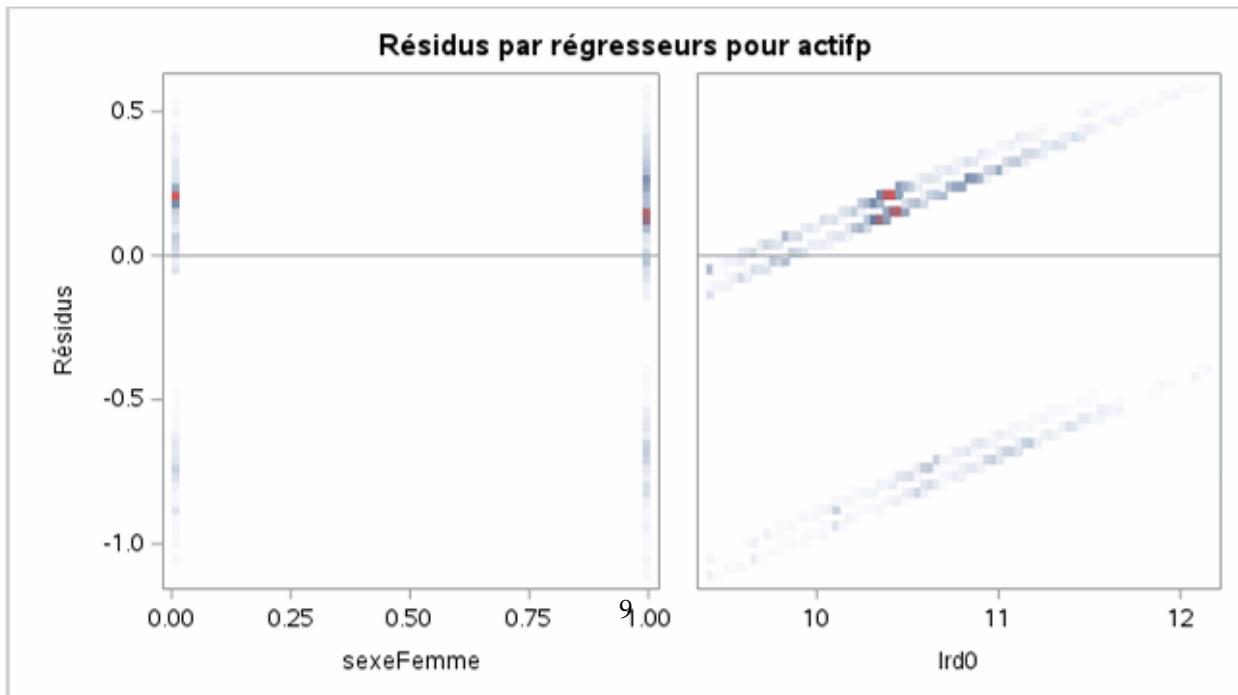
Analyse de variance					
Source	DDL	Somme des carrés	Moyenne quadratique	Valeur F	Pr > F
Modèle	2	154.54008	77.27004	528.90	<.0001
Erreur	10545	1540.56534	0.14609		
Total sommes corrigées	10547	1695.10542			

Root MSE	0.38222	R carré	0.0912
Moyenne dépendante	0.79882	R car. ajust.	0.0910
Coeff Var	47.84818		

Paramètres estimés					
Variable	DDL	Valeur estimée des paramètres	Erreur type	Valeur du test t	Pr > t
Intercept	1	3.55112	0.08479	41.88	<.0001
sexefemme	1	0.06835	0.00793	8.62	<.0001
lrd0	1	-0.26606	0.00818	-32.52	<.0001

Régression de actifp sur sexeFemme et lrd0

La procédure REG
Modèle : MODEL1
Variable dépendante : actifp



Régression de actifs sur sexeFemme et lrd0

La procédure REG

Modèle : MODEL1

Variable dépendante : actifs

Nb d'observations lues	10548
Nb d'obs. utilisées	10548

Analyse de variance					
Source	DDL	Somme des carrés	Moyenne quadratique	Valeur F	Pr > F
Modèle	2	2.97787	1.48894	49.14	<.0001
Erreur	10545	319.50933	0.03030		
Total sommes corrigées	10547	322.48720			

Root MSE	0.17407	R carré	0.0092
Moyenne dépendante	0.03157	R car. ajust.	0.0090
Coeff Var	551.37157		

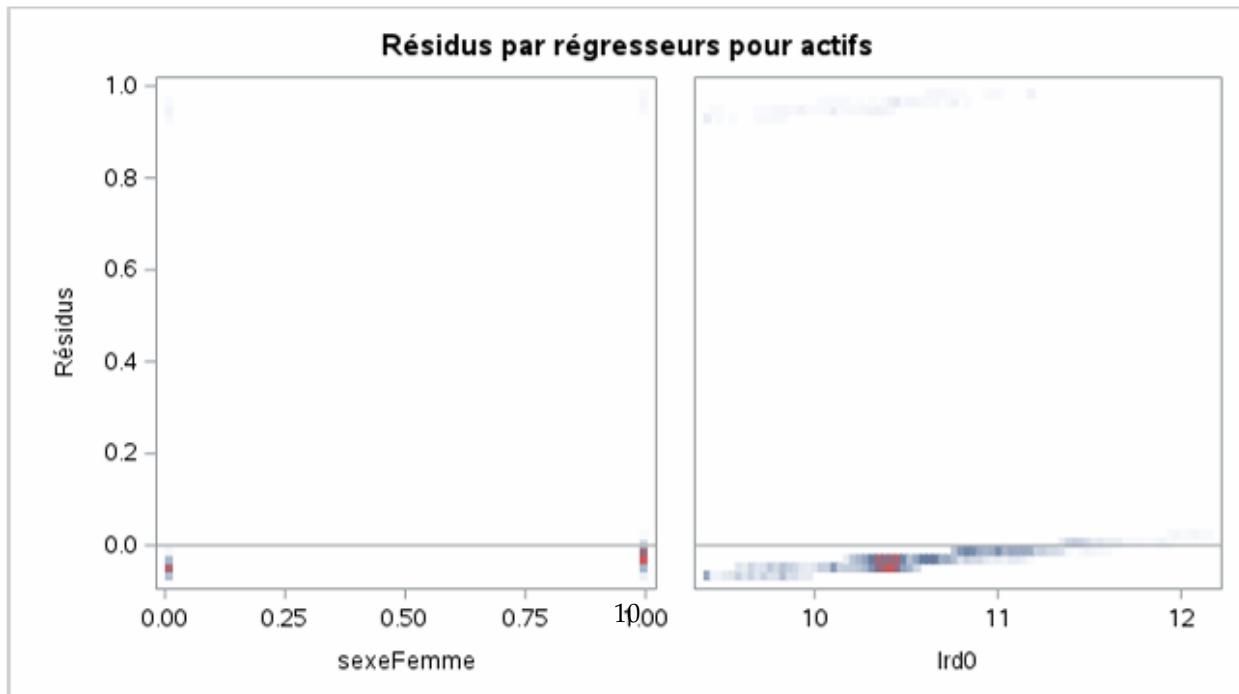
Paramètres estimés					
Variable	DDL	Valeur estimée des paramètres	Erreur type	Valeur du test t	Pr > t
Intercept	1	0.35447	0.03862	9.18	<.0001
sexxFemme	1	-0.01201	0.00361	-3.33	0.0009
lrd0	1	-0.03006	0.00373	-8.07	<.0001

Régression de actifs sur sexeFemme et lrd0

La procédure REG

Modèle : MODEL1

Variable dépendante : actifs



Premièrement, nous pouvons lire que les modèles visant à expliquer l'activité principale ainsi que l'activité secondaire possèdent un coefficient de détermination d'environ 0,9%, soit la capacité des variables explicatives à expliquer les variances de ces modèles, ce qui est plutôt faible.

L'ensemble des coefficients sont statistiquement significatifs individuellement d'après le test de student et notons une significativité globale selon les résultats du test de Fisher. De manière générale, nous observons que le fait d'être un homme à effet négatif sur la probabilité de réalisation de 'actifp'. En effet un femme aura alors 6,8% de chance qu'un homme de réaliser une activité principale (*cf. modèle "Régression de actifp sur sexeFemme et lrd0"*), *ceteris paribus*. Ces modèles nous montrent également que le fait d'être un homme augmente la probabilité que 'actifs' se réalise. En effet, un homme aura alors 1% de chance de plus qu'un homme de réaliser une activité secondaire (*cf. modèle "Régression de actifs sur sexeHomme et lrd0"*), *ceteris paribus*.

Concernant les modèles "Régression de actifp sur sexeN et lrd0" et "Régression de actifs sur sexeN et lrd0", bien que l'on puisse observer un signe négatif lorsque l'individu est un homme, de même manière que dans l'exercice précédent, la constante compense cet effet par rapport au modèle avec un sexe représenté par une indicatrice binaire. En effet, il est logique de remarquer qu les effets obtenus dans les modèles contenant la variable sexeN seront identiques que ceux des modèles contenant les variables sexeHomme/sexeFemme, seuls l'interprétation de ces modèles sera différente du fait que la spécification soit différente.

Notons par ailleurs, que le coefficient du revenu des individus inactifs sera identique pour tous les modèles expliquant l'activité principale (*cf. Régression de actifp sur sexeN/sexeHomme/sexeFemme et lrd0*) et tous les modèles expliquant l'activité secondaire (*cf. Régression de actifs sur sexeN/sexeHomme/sexeFemme et lrd0*). En effet, en moyenne, une augmentation de revenu 1% pour un individu inactif fera diminuer les chances d'avoir une activité principale de 0,26% et une activité secondaire de 0,03%, *ceteris paribus*. Cependant, il pourrait être intéressant d'étendre l'analyse en incluant des termes d'interaction entre le sexe et le revenu d'inactivité afin de pouvoir observer s'il y existe des différences de cet effet en fonction du sexe ou encore d'inclure un terme quadratique afin d'observer s'il existe un effet de seuil sur l'activité concernant le revenu d'inactivité.

3 Question 3 : Reprendre la question avec un modèle logit puis avec un modèle Probit. Commenter les similitudes et différences par rapport aux MCO.

```
----- QUESTION 3-LOGIT-----*/  
/*Régression actifp et actifs expliqué par sexeN */  
  
proc logistic data=merged_panelex3;  
    class sexe;  
    model actifp = sexeN lrd0 / link=logit;  
    title "Modèle logit de actifp sur sexeN et lrd0";  
run;  
quit;  
  
proc logistic data=merged_panelex3;  
    class sexe;  
    model actifs = sexeN lrd0 / link=logit;  
    title "Modèle logit de actifs sur sexeN et lrd0";  
run;  
quit;  
  
/*Régression actifp et actifs expliqué par sexeHomme */  
proc logistic data=merged_panelex3;  
    class sexe;  
    model actifp = sexeHomme lrd0 / link=logit;  
    title "Modèle logit de actifp sur sexeHomme et lrd0";  
run;  
quit;  
  
proc logistic data=merged_panelex3;  
    class sexe;  
    model actifs = sexeHomme lrd0 / link=logit;  
    title "Modèle logit de actifs sur sexeHomme et lrd0";  
run;  
quit;  
  
/*Régression actifp et actifs expliqué par sexeFemme */  
proc logistic data=merged_panelex3;
```

```

class sexe;
model actifp = sexeFemme lrd0 / link=logit;
title "Modèle logit de actifp sur sexeFemme et lrd0";
run;
quit;

proc logistic data=merged_panelex3;
class sexe;
model actifs = sexeFemme lrd0 / link=logit;
title "Modèle logit de actifs sur sexeFemme et lrd0";
run;
quit;

```

Figure 1. Modèle logit de actifp sur sexeN et lrd0

Analyse des valeurs estimées du maximum de vraisemblance					
Paramètre	DDL	Estimation	Erreur type	Khi-2 de Wald	Pr > khi-2
Intercept	1	20.4867	0.6600	963.5245	<.0001
sexen	1	0.5785	0.0570	103.1682	<.0001
lrd0	1	-1.8895	0.0646	854.8933	<.0001

Figure 2. Modèle logit de actifs sur sexeN et lrd0

Analyse des valeurs estimées du maximum de vraisemblance					
Paramètre	DDL	Estimation	Erreur type	Khi-2 de Wald	Pr > khi-2
Intercept	1	8.0487	1.2748	39.8612	<.0001
sexen	1	-0.4072	0.1143	12.7005	0.0004
lrd0	1	-1.0447	0.1260	68.7976	<.0001

Figure 3. Modèle logit de actifp sur sexeHomme et lrd0

Analyse des valeurs estimées du maximum de vraisemblance					
Paramètre	DDL	Estimation	Erreur type	Khi-2 de Wald	Pr > khi-2
Intercept	1	21.6436	0.7005	954.6542	<.0001
sexehomme	1	-0.5785	0.0570	103.1682	<.0001
lrd0	1	-1.8895	0.0646	854.8933	<.0001

Figure 4. Modèle logit de actifs sur sexeHomme et lrd0

Analyse des valeurs estimées du maximum de vraisemblance					
Paramètre	DDL	Estimation	Erreur type	Khi-2 de Wald	Pr > khi-2
Intercept	1	7.2343	1.3078	30.5986	<.0001
sexehomme	1	0.4072	0.1143	12.7005	0.0004
lrd0	1	-1.0447	0.1260	68.7976	<.0001

Figure 5. Modèle logit de actifp sur sexeFemme et lrd0

Analyse des valeurs estimées du maximum de vraisemblance					
Paramètre	DDL	Estimation	Erreur type	Khi-2 de Wald	Pr > khi-2
Intercept	1	21.0652	0.6782	964.8598	<.0001
sexeFemme	1	0.5785	0.0570	103.1682	<.0001
lrd0	1	-1.8895	0.0646	854.8933	<.0001

Figure 6. Modèle logit de actifs sur sexeFemme et lrd0

Analyse des valeurs estimées du maximum de vraisemblance					
Paramètre	DDL	Estimation	Erreur type	Khi-2 de Wald	Pr > khi-2
Intercept	1	7.6415	1.2864	35.2884	<.0001
sexeFemme	1	-0.4072	0.1143	12.7005	0.0004
lrd0	1	-1.0447	0.1260	68.7976	<.0001

```

/*----- QUESTION 3-PROBIT-----*/
/*Régression actifp et actifs expliqué par sexeN */

proc logistic data=merged_panelex3;
  class sexe;
  model actifp = sexeN lrd0 / link=probit;
  title "Modèle probit de actifp sur sexeN et lrd0";
run;
quit;

proc logistic data=merged_panelex3;
  class sexe;
  model actifs = sexeN lrd0 / link=probit;
  title "Modèle probit de actifs sur sexeN et lrd0";
run;
quit;

/*Régression actifp et actifs expliqué par sexeHomme */

proc logistic data=merged_panelex3;
  class sexe;
  model actifp = sexeHomme lrd0 / link=probit;
  title "Régression de actifp sur sexeHomme et lrd0";
run;
quit;

proc logistic data=merged_panelex3;
  class sexe;
  model actifs = sexeHomme lrd0 / link=probit;
  title "Modèle probit de actifs sur sexeHomme et lrd0";
run;
quit;

/*Régression actifp et actifs expliqué par sexeFemme */

proc logistic data=merged_panelex3;
  model actifp = sexeFemme lrd0 / link=probit;
  title "Modèle probit de actifp sur sexeFemme et lrd0";
run;
quit;

proc logistic data=merged_panelex3;
  class sexe;

```

```

model actifs = sexeFemme lrd0 / link=probit;
title "Modèle probit de actifs sur sexeFemme et lrd0";
run;
quit;

```

Figure 7. Modèle probit de actifp sur sexeN et lrd0

Analyse des valeurs estimées du maximum de vraisemblance					
Paramètre	DDL	Estimation	Erreur type	Khi-2 de Wald	Pr > khi-2
Intercept	1	-11.2746	0.3620	970.2633	<.0001
sexen	1	-0.3027	0.0318	90.3882	<.0001
lrd0	1	1.0320	0.0355	846.3369	<.0001

Figure 8. Modèle probit de actifs sur sexeN et lrd0

Analyse des valeurs estimées du maximum de vraisemblance					
Paramètre	DDL	Estimation	Erreur type	Khi-2 de Wald	Pr > khi-2
Intercept	1	3.2471	0.5662	32.8894	<.0001
sexen	1	-0.1693	0.0500	11.4664	0.0007
lrd0	1	-0.4659	0.0557	69.9511	<.0001

Figure 9. Modèle probit de actifp sur sexeHomme et lrd0

Analyse des valeurs estimées du maximum de vraisemblance					
Paramètre	DDL	Estimation	Erreur type	Khi-2 de Wald	Pr > khi-2
Intercept	1	11.8800	0.3821	966.8057	<.0001
sexehomme	1	-0.3027	0.0318	90.3882	<.0001
lrd0	1	-1.0320	0.0355	846.3369	<.0001

Figure 10. Modèle probit de actifs sur sexeHomme et lrd0

Analyse des valeurs estimées du maximum de vraisemblance					
Paramètre	DDL	Estimation	Erreur type	Khi-2 de Wald	Pr > khi-2
Intercept	1	2.9085	0.5813	25.0355	<.0001
sexehomme	1	0.1693	0.0500	11.4664	0.0007
lrd0	1	-0.4659	0.0557	69.9511	<.0001

Figure 11. Modèle probit de actifp sur sexeFemme et lrd0

Analyse des valeurs estimées du maximum de vraisemblance					
Paramètre	DDL	Estimation	Erreur type	Khi-2 de Wald	Pr > khi-2
Intercept	1	11.5773	0.3708	974.9136	<.0001
sexefemme	1	0.3027	0.0318	90.3882	<.0001
lrd0	1	-1.0320	0.0355	846.3369	<.0001

Figure 12. Modèle probit de actifs sur sexeFemme et lrd0

Analyse des valeurs estimées du maximum de vraisemblance					
Paramètre	DDL	Estimation	Erreur type	Khi-2 de Wald	Pr > khi-2
Intercept	1	3.0778	0.5716	28.9922	<.0001
sexefemme	1	-0.1693	0.0500	11.4664	0.0007
lrd0	1	-0.4659	0.0557	69.9511	<.0001

Tout d'abord, nous remarquons que l'ensemble des estimations de coefficients sont significatif selon les tests du Khi-2 de Wald.

Bien que seuls le signe des estimations soit interprétable dans nos modèles logit et probit, nous remarquons que pour l'ensemble des modèles logit, le coefficient estimé du logarithme du revenu d'inactivité sera de -1,88 pour les modèles prédisant l'activité principale et de -1,04 pour les modèles prédisant l'activité secondaire. De même pour l'ensemble des modèles probit, pour lesquels, le coefficient estimé de lrd0 sera de -1,03 pour les modèles prédisant actifp et de -0,46 pour les modèles prédisant actifs. Ces résultats suggèrent qu'il existe bien quatre groupes de trois modèles identiques.

Concernant les prédictions, tout comme dans nos modèles MCO, le signe de lrd0 révèle dans chacun des modèles que ce dernier diminue les chances que actifp/actifs se réalisent, ceteris paribus. Par ailleurs,

nous remarquons que la probabilité que l'activité principale se réalise lorsque l'individu est un femme augmente, tandis que lorsque l'individu est un homme la probabilité de la réalisation de l'activité secondaire augmente, *ceteris paribus*.

De ce fait, bien que l'on ne puisse pas interpréter la magnitude des coefficients estimés dans ce modèle, nous pouvons voir que les effets (positifs et négatifs) sur nos variables actifp et actifs seront identiques dans nos modèles MCO, logit et probit.

4 Question 4 : (papier-crayon)

①

$$\text{actif}_P = \beta_0 + \beta_1 \text{SexeFemme} + \beta_2 \text{LrdO}$$

(ou sexeHomme)

$$\text{actif}_S = \beta_0 + \beta_1 \text{SexeFemme} + \beta_2 \text{LrdO}$$

(ou sexeHomme)

Logit (actif_P):

Lorsque l'individu est une Femme : un Homme:

$$\text{actif}_P = 21,06 + 0,57 - 1,88 \text{LrdO}$$

$$\text{actif}_P = 21,06 - 1,88 \text{LrdO}$$

Logit (actif_S):

Lorsque l'individu est une femme : un homme:

$$\text{actif}_S = 7,64 - 0,40 - 1,04 \text{LrdO}$$

$$\text{actif}_S = 7,64 - 1,04 \text{LrdO}$$

Pour un modèle Logit nous avons une fonction de répartition telle que :

$$F(u) = \frac{1}{1 + \exp(-u)}$$

Ainsi pour un LrdO = 3,5, nous aurons pour une femme :

$$\text{actif}_P = 21,0652 + 0,5785 - 1,8855$$

$$\text{actif}_P = 3,69$$

$$\text{Ainsi, } F(3,69) = \frac{1}{1 + \exp(-3,69)} = 0,97$$

Table de probabilité d'activité principale en fonction de LrdO pour un modèle Logit:

LrdO	Femme	Homme	Eff. marg. Femme	Eff. marg. homme
3,5	0,97	0,95	-0,04	-0,07
10	0,93	0,89	-0,10	-0,17
11	0,70	0,56	-0,39	-0,46
11,5	0,47	0,33	-0,97	-0,42
12	0,26	0,16	-0,36	-0,26

Pour LrdO, nous obtenons l'effet marginal pour une femme avec : $\beta_2 \times P(1 - P)$

$$= 1,8895 \times 0,976(1 - 0,976)$$

(2)

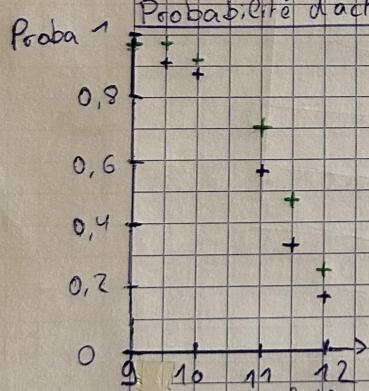
Table de probabilité d'activité secondaire en fonction de l'rd0 pour un modèle logit:

l'rd0	Femme	homme	Eff. marg. femme	Eff. marg. homme
9,5	0,06	0,09	-0,03	-0,09
10	0,03	0,05	-0,03	-0,04
11	0,01	0,02	-0,01	-0,02
11,5	0,008	0,01	-0,008	-0,01
12	0,005	0,007	-0,005	-0,007

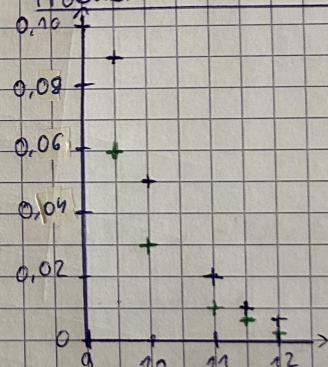
Le signe des effets marginaux révèlent une fonction décroissante

Représentations graphiques (logit):

Probabilité d'activité p.



Probabilité d'activité s:



- homme
- femme

(3)

Probit (actif; Fp):

Lorsque l'individu est une femme : un homme :

$$\text{actif}_F = 11,57 + 0,30 - 1,03 \text{ lrd}0 \quad \text{actif}_P = 11,57 - 1,03 \text{ lrd}0$$

Probit (actifs):

Lorsque l'individu est une femme : un homme :

$$\text{actif}_S = 3,07 - 0,16 - 0,46 \text{ lrd}0 \quad \text{actif}_S = 3,07 - 0,46 \text{ lrd}0$$

Pour un modèle probit, nous avons une fonction de répartition telle que :

$$F(u) = \int_{-\infty}^u \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) dx$$

Ainsi pour un $\text{lrd}0 = 9,5$, nous aurons pour une femme :

$$\text{actif}_F = 11,57 + 0,3027 - 1,0320 \approx 9,5$$

$$\text{actif}_P = 2,076$$

Pour $u = \text{actif}_P$:

$$F(u) = 0,981$$

L'effet marginal s'obtenant par : $F'_x \times P = 1,032 \times 0,981$

$$= 1,01$$

Table de probabilité d'activité principale en fonction de lrd0 pour un modèle probit:

lrd0	Femme	homme	Eff. mary. femme	Eff. mary. homme
9,5	0,98	0,96	-1,01	-0,99
10	0,94	0,89	-0,97	-0,92
11	0,70	0,58	-0,72	-0,60
11,5	0,50	0,38	-0,52	-0,39
12	0,30	0,20	-0,31	-0,21

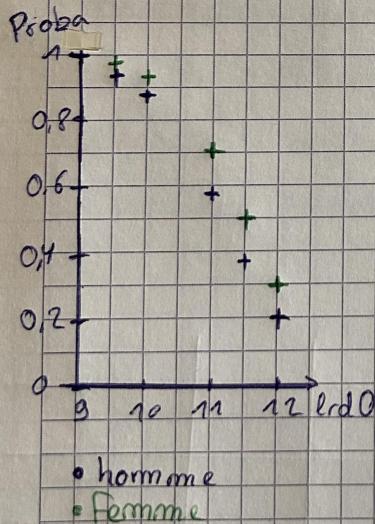
(4)

Table de probabilité d'activité secondaire en fonction de l'rd0 pour un modèle probit:

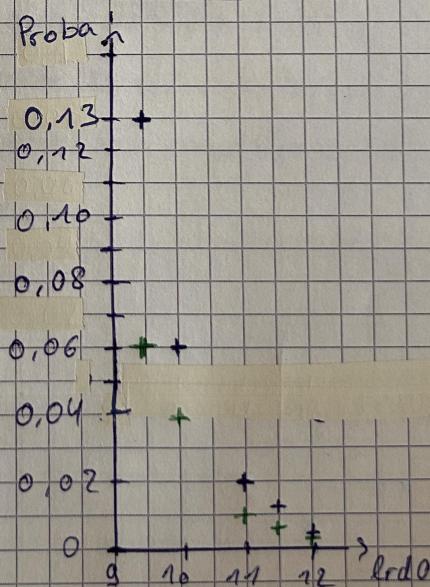
lrd0	Femme	homme	Eff. marg. Femme	Eff. marg. homme
9,5	0,06	0,13	-0,2	-0,06
10	0,04	0,06	-0,01	-0,02
11	0,01	0,02	-0,006	-0,009
11,5	0,007	0,011	-0,003	-0,005
12	0,004	0,006	-0,001	-0,002

Représentations graphiques (probit):

Probabilité de réalisation
de l'activité en fonction
du sexe:



Probabilité de réalisation
de l'activité en fonction
du sexe:



⑤

Table des différences de probabilité
en fonction du sexe et du revenu d'inactivité
(en valeur absolue).

erdo	Logit-actifp	Logit-actifs	Probit-actifp	Probit-actifs
9,5	0,02	0,03	0,02	0,07
10	0,04	0,02	0,05	0,02
11	0,14	0,01	0,12	0,01
11,5	0,14	0,002	0,12	0,004
12	0,10	0,002	0,10	0,002

La différence dans l'intervalle $[9,5; 12]$ semble s'agrandir pour les modèles prédisant l'activité principale avant de diminuer (pour $\text{lod}\theta=12$) tandis que la différence semble se réduire lorsque le revenu d'inactivité augmente pour les modèles prédisant l'activité secondaire.

5 Question 5 : Calculer ces différentes probabilités pour chaque ligne de l'échantillon et tracer sur un même graphique toutes les courbes de probabilité d'activité principale en fonction du log-revenu d'inactivité. Idem pour la probabilité d'activité secondaire. Vérifier la cohérence avec la question précédente.

```
----- QUESTION 5-LOGIT -----
proc logistic data=merged_panelex2;
    model actifp(event='1') = lrd0 sexeFemme / link=logit;
    output out=output_actifp1 predicted=pred_actifp1;
run;

proc sgplot data=output_actifp1;
    series x=lrd0 y=pred_actifp1 / markers group=sexeFemme;
    yaxis label="Probabilité de réalisation de l'activité principale";
    xaxis label="Logarithme du revenu d'inactivité";
    keylegend / title="Sexe";
run;

proc logistic data=merged_panelex2;
    model actifs(event='1') = lrd0 sexeFemme / link=logit;
    output out=output_actifs1 predicted=pred_actifs1;
run;

proc sgplot data=output_actifs1;
    series x=lrd0 y=pred_actifs1 / markers group=sexeFemme;
    yaxis label="Probabilité de réalisation de l'activité secondaire";
    xaxis label="Logarithme du revenu d'inactivité";
    keylegend / title="Sexe";
run;
```

Figure 13. Probabilité de réalisation de l'activité principale en fonction du revenu d'inactivité et du sexe - Modèle Probit

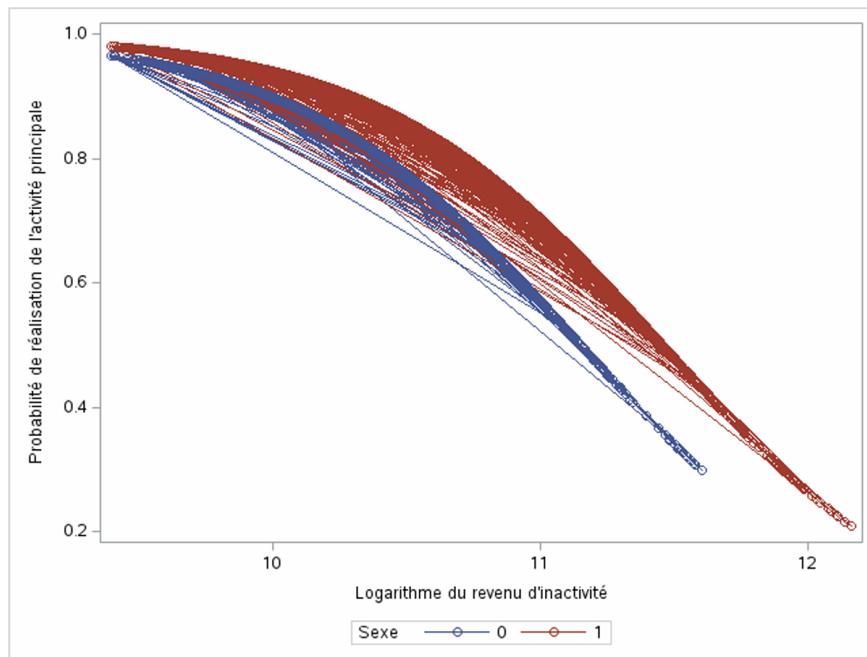
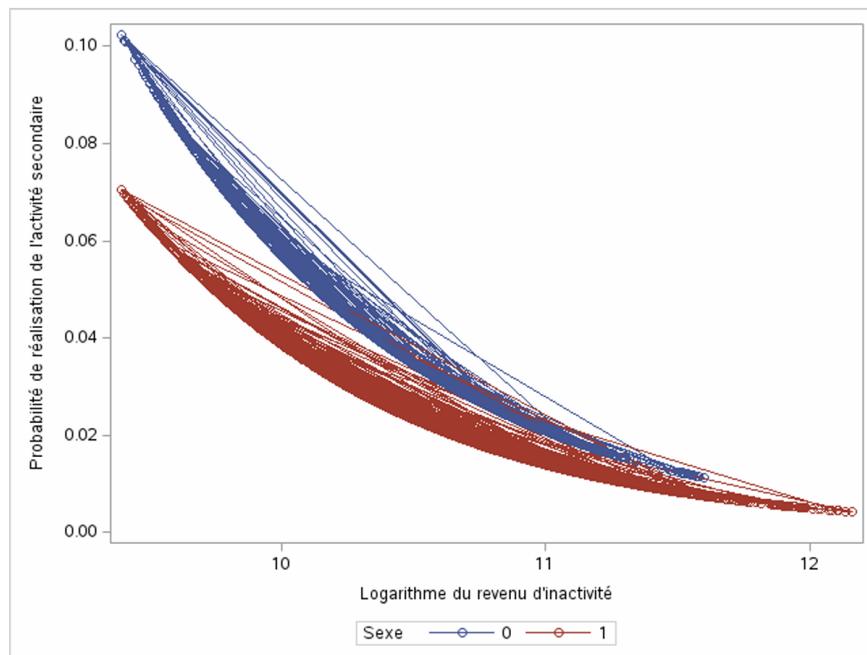


Figure 14. Probabilité de réalisation de l'activité secondaire en fonction du revenu d'inactivité et du sexe - Modèle Probit



```
/*----- QUESTION 5-PROBIT-----*/  
proc logistic data=merged_panelex2;  
    model actifp(event='1') = lrd0 sexeFemme / link=probit;  
    output out=output_actifp2 predicted=pred_actifp2;  
run;  
  
proc sgplot data=output_actifp2;  
    series x=lrd0 y=pred_actifp2 / markers group=sexeFemme;  
    yaxis label="Probabilité de réalisation de l'activité principale";  
    xaxis label="Logarithme du revenu d'inaktivité";  
    keylegend / title="Sexe";  
run;  
  
proc logistic data=merged_panelex2;  
    model actifs(event='1') = lrd0 sexeFemme/ link=probit;  
    output out=output_actifs2 predicted=pred_actifs2;  
run;  
  
proc sgplot data=output_actifs2;  
    series x=lrd0 y=pred_actifs2 / markers group=sexeFemme;  
    yaxis label="Probabilité de réalisation de l'activité secondaire";  
    xaxis label="Logarithme du revenu d'inaktivité";  
    keylegend / title="Sexe";  
run;
```

Figure 15. Probabilité de réalisation de l'activité principale en fonction du revenu d'inactivité et du sexe - Modèle Logit

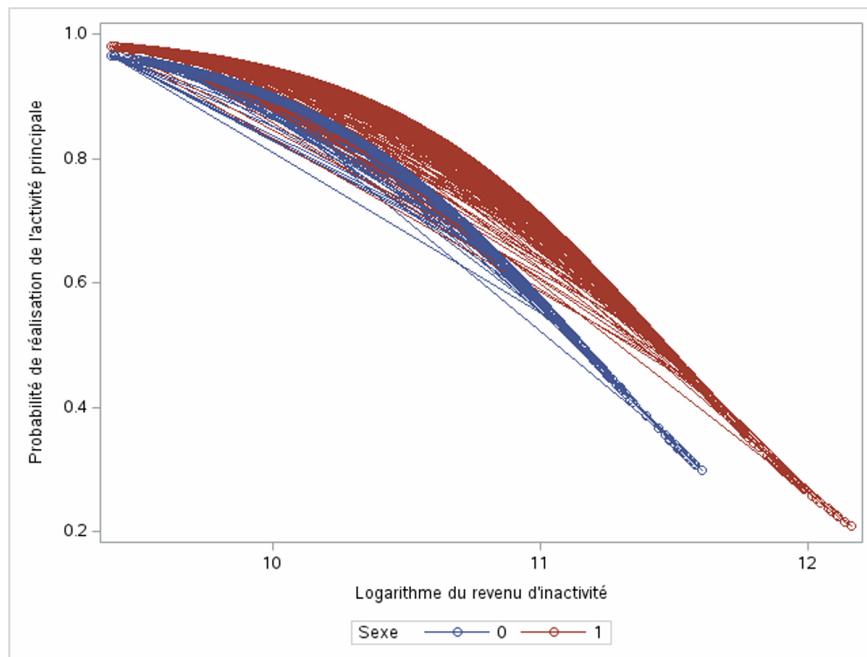
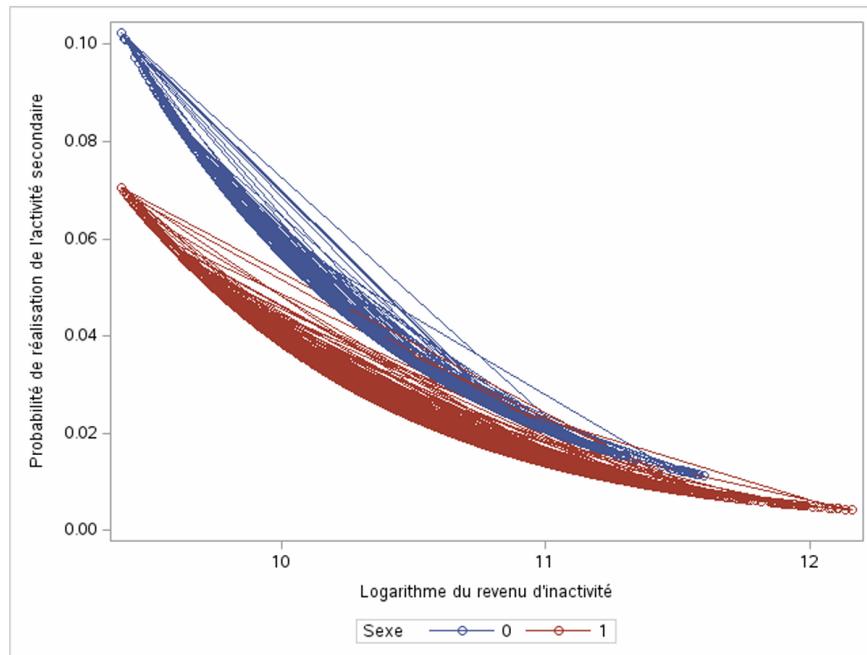


Figure 16. Probabilité de réalisation de l'activité secondaire en fonction du revenu d'inactivité et du sexe - Modèle Logit



Ces graphiques nous permettent d'observer visuellement nos interprétations précédentes. Nous pouvons observer une allure décroissante et non linéaire pour les courbes représentant la relation entre le logarithme du revenu d'inactivité et la probabilité de réalisation d'une activité principale et secondaire. Concernant les modèles prédisant la réalisation de l'activité principale, nous observons une relation concave, traduisant une probabilité élevée de la réalisation de l'événement jusqu'à un point d'inflection (environ 10,5) à partir duquel les chances de réalisation vont devenir de plus en plus décroissante tandis que pour les modèles prédisant la réalisation de l'activité secondaire, nous observons une relation convexe, traduisant une diminution du taux de décroissance des probabilités de réalisation de l'événement à partir d'un point d'inflection (environ 10,5).

Par ailleurs, dans nos graphiques représentant les prédictions de la réalisation de actifp nous montre une courbe rouge représentant l'observation des individus de sexe féminin est au dessus de la courbe bleue, témoignant des chances supplémentaires pour une femme de réaliser une activité principale par rapport à un homme. Cette interprétation sera identique pour les individus de sexe masculin concernant la probabilité de réalisation d'activité secondaire.

De plus, nous pouvons également interpréter la distance entre les courbes représentant les hommes et femmes comme outil de comparaison. De ce fait, il est intéressant d'observer premièrement que dans les modèles prédisant l'activité principale, qu'à mesure le logarithme du revenu d'inactivité diminue, les hommes et les femmes semblent arriver à une même probabilité élevée de réalisation de l'activité secondaire tandis qu'à mesure que le revenu d'inactivité augmente, la probabilité de réalisation de l'activité principale semble se différencier en fonction du sexe. Deuxièmement, nous observons que pour les modèles prédisant l'activité secondaire, qu'à mesure où le revenu d'inactivité diminue, il y a convergence des probabilités vers de faibles chances de réalisation de l'événement.

6 Question 6 :

Dans cet exercice, nous imposons une contrainte à "lrd0" :

- Pour les hommes : lrd0 reste lrd0
- Pour les Femmes : lrd0 devient $lrd0 - \frac{\beta_2^m}{\beta_2^m}$

Désormais, avec m pouvant signifier MCO, logit ou probit :

$$E(\text{actifp} | \text{SexeFemme} = 0) = \beta_0 + \beta_2^m lrd0$$

$$\begin{aligned} E(\text{actifp} | \text{SexeFemme} = 1) &= \beta_0 + \beta_1^m + \beta_2^m \left[lrd0 - \frac{\beta_2^m}{\beta_2^m} \right] \\ &= \beta_0 + \beta_1^m + \beta_2^m lrd0 - \beta_2^m \frac{\beta_2^m}{\beta_2^m} \\ &= \beta_0 + \beta_1^m + \beta_2^m lrd0 - \beta_2^m \end{aligned}$$

$$E(\text{actifp} | \text{SexeFemme} = 1) = \beta_0 + \beta_2^m lrd0$$

Ainsi nous remarquons que :

$$E(\text{actifp} | \text{SexeFemme} = 0) = E(\text{actifp} | \text{SexeFemme} = 1)$$

Le modèle étant identique pour l'explication de l'activité secondaire nous obtiendrons la même limite dû à l'instauration d'une variable contrainte, Vm.

$$E(\text{actifs} | \text{SexeFemme} = 0) = E(\text{actifs} | \text{SexeFemme} = 1)$$

De ce fait, notre modèle ne considère plus la différence de sexe, révélant une complexité pour les modèles binaires à traiter les interactions.