

Micro-simuler l'impôt sur le revenu

1 Introduction

Je me propose de simuler le calcul de l'impôt sur le revenu à partir des données de l'enquête *emploi* de l'INSEE. Les revenus des ménages sont toutefois pas bien connus dans cette enquête dont l'objet est d'abord de caractériser les forces de travail en FRANCE. En outre, nous allons nous restreindre aux ménages qui forment un seul foyer fiscal en supposant que l'éventuel couple dans le ménage est un couple de personnes mariées ou pacsées. Comme cela, nous n'avons pas à ventiler les ménages en plusieurs foyers fiscaux.

Il va donc nous falloir retenir dans un premier temps une sous population particulière dans l'enquête *emploi*. Dans un second temps, nous allons programmer sous SAS le calcul d'une forme très simplifiée de l'impôt sur le revenu.

2 La sélection de la population d'intérêt

En fait, dans l'enquête *emploi*, on ne dispose que de la variable `salred` (pour *salaires redressés*) pour les actifs occupés qui sont salariés. Je me propose ainsi de retenir les ménages pour lesquels il n'y a pas d'incertitudes sur les revenus, c'est-à-dire *grossomodo* les ménages dont la personne de référence est en âge de travailler et les membres sont soit salariés soit inactifs.

Comme les données sont particulièrement volumineuses, dans un premier temps, je sélectionne les variables d'intérêt qui sont les suivantes.

1. `ident` — l'identifiant du ménage ;
2. `lpr` — le lien avec la personne de référence du ménage ;
3. `salred` — le salaire mensuel redressé ;
4. `ag` — l'âge au moment de l'enquête.

En outre, je ne prends que les observations relatives au quatrième trimestre de l'année. Le fragment du programme SAS est ainsi le suivant.

```
1 DATA brutes ;
2   SET ee.base_entiere ;
3   IF trim EQ 4 ;
4   KEEP ident lpr salred act ag ;
```

```
5 RUN ;
6 PROC sort DATA = brutes OUT = brutes ;
7   BY ident ;
8 RUN ;
9 DATA compter_menages (KEEP = ident) ;
10  SET brutes ;
11  BY ident ;
12  IF FIRST.ident THEN OUTPUT ;
13 RUN ;
```

La dernière étape DATA est destinée à compter le nombre initial de ménages ; je n'ai pas trouvé une manière plus simple de le faire.

Pour sélectionner les ménages d'intérêt, le plus simple est de se définir un drapeau, initialement abaissé, que l'on va lever dès que l'on trouve un membre du ménage qui ne satisfait pas les conditions que l'on s'est fixées. On cherche à ce que

1. le ménage soit constitué d'un seul foyer fiscal et donc que le lien avec la personne de référence du ménage soit
 - personne de référence,
 - conjoint de la personne de référence,
 - enfant de la personne de référence ;
2. l'âge de la personne de référence du ménage soit compris entre 20 et 60 ans ;
3. les actifs occupés disposent d'un salaire redressé renseigné, c'est-à-dire strictement positif ;
4. les chômeurs soient exclus puisque l'on ne connaît pas leurs revenus.

Une première étape DATA est destinée à repérer les ménages que l'on ne veut pas retenir. La seconde étape va permettre de ne garder que les individus des ménages que l'on veut retenir.

Le fragment de programme est finalement le suivant.

```
1 %LET lpr_pr = '1' ;
2 %LET lpr_conj = '2' ;
3 %LET lpr_enf = '3' ;
4 %LET act_occ = '1' ;
5 %LET act_inactif = '3' ;
6 DATA pas_bons (KEEP = ident) ;
7   SET brutes ;
8   BY ident ;
9   RETAIN flag ;
10  IF FIRST.ident THEN flag = 0 ;
11  IF (lpr EQ &lpr_pr) AND NOT (20 <= ag <= 60) THEN flag = 1 ;
12  IF NOT lpr IN(&lpr_pr, &lpr_conj, &lpr_enf) THEN flag = 1 ;
13  IF (lpr EQ &lpr_pr) AND NOT ((act EQ &act_occ AND salred > 0)
```

```

14 OR (act EQ &act_inactif)) THEN flag = 1 ;
15 IF (lpr EQ &lpr_conj) AND NOT ((act EQ &act_occ AND salred > 0)
16 OR (act EQ &act_inactif)) THEN flag = 1 ;
17 IF LAST.ident AND flag THEN OUTPUT ;
18 RUN ;
19 DATA foyers ;
20 MERGE brutes pas_bons (IN = pas_bons) ;
21 BY ident ;
22 IF NOT pas_bons ;
23 RUN ;

```

Notez l'utilisation du macro-processeur pour spécifier les différentes valeurs de la variable lpr ou de la variable act ; c'est important de rendre le programme expressif.

Les ménages se répartissent de la sorte comme suit.

Ménages	Nombre
Initialement dans l'enquête <i>emploi</i>	36 622
À ne pas retenir	28 704
À retenir	7 918

3 Les paramètres de la législation fiscale

C'est une bonne idée, me semble-t-il, de stocker les paramètres de la législation fiscale dans une table Sas. Cette table ne comporte qu'une seule observation et le nom des variables est le nom que l'on donne au paramètre. Le plus simple est d'utiliser une variable pour nommer les paramètres et d'utiliser ensuite l'option ID de la PROC transpose. Le fragment de programme est ainsi le suivant.

```

1 DATA prm ;
2 INPUT nom $ valeur ;
3 CARDS ;
4 tauxfp .1
5 minfp 401
6 maxfp 13501
7 taux0 0
8 seuil0 0
9 taux1 5.5
10 seuil1 5687
11 taux2 14
12 seuil2 11344
13 taux3 30
14 seuil3 25195
15 taux4 40
16 seuil4 67546

```

```

17 RUN ;
18 PROC transpose DATA = prm OUT = prm ;
19 ID nom ;
20 RUN ;

```

4 Le programme de simulation de l'impôt sur le revenu

Le programme de simulation de l'impôt sur le revenu figure ci-après. Il est outrageusement simplifié :

1. les demi-parts de quotient familial pour les foyers mono-parentaux ne sont pas prises en compte ;
2. le plafonnement de l'avantage en impôt apporté par les demi-parts des enfants n'est pas programmé ;
3. de nombreux aspects de la législation qui pourraient être retracés en mobilisant l'information apportée par l'enquête *emploi* ne sont pas modélisés.

```

1 DATA impot ;
2 IF _N_ EQ 1 THEN SET prm ;
3 ARRAY taux taux0 - taux4 ;
4 ARRAY seuils seuil0 - seuil4 ;
5 SET foyers ;
6 BY ident ;
7 RETAIN assiette nbadultes nbenfants ;
8 IF FIRST.ident THEN DO ;
9 assiette = 0 ;
10 nbadultes = 0 ;
11 nbenfants = 0 ;
12 END ;
13 IF salred > 0 THEN DO ;
14 fp = MIN(MAX(minfp, tauxfp*salred*12), maxfp) ;
15 assiette = assiette + MAX(0, salred*12 - fp) ;
16 END ;
17 IF lpr IN(&lpr_pr, &lpr_conj) THEN nbadultes = nbadultes + 1 ;
18 IF lpr EQ &lpr_enf THEN nbenfants = nbenfants + 1 ;
19 IF LAST.ident THEN DO ;
20 nbparts = nbadultes + .5*nbenfants ;
21 r = assiette / nbparts ;
22 imp = 0 ;
23 DO i = 2 TO 5 ;
24 imp = imp + (taux{i}-taux{i-1})/100 * max(0, r-seuils{i}) ;
25 END ;
26 imp = nbparts*imp ;

```

```
27  OUTPUT ;  
28  KEEP imp assiette nbadultes nbenfants ;  
29  END ;  
30  RUN ;
```