

# Fiche de modélisations n°3

Reproduction des résultats de l'article de Duvoux et Papuchon

Kim Antunez

## 1 Objectif

L'objectif de cette troisième série de modèles est de reproduire les résultats statistiques contenus dans l'article de Duvoux et Papuchon afin de s'assurer de bien **comprendre leur démarche empirique**.

## 2 Analyses

Pas grand chose de plus que ce qui est contenu dans leur article. Les résultats de l'article ont pu être reproduits avec parfois des écarts d'un ou deux points de pourcentages sûrement dûs à des petites différences de champs pour certaines variables. Différentes remarques ont été ajoutées au fil du code.

## 3 Code et résultats

```
#chargement des packages
library(knitr)
library(dplyr) #manipuler les bases de données
library(tidyr) #pour pivot_longer
library(ggplot2) #pour les graphiques
```

```
#Chargement de la base de données
bdd <- readRDS("../data/2017/barometre2000_2017_papuchon.rds") %>%
  filter(annee%in%2015:2017)
```

### 3.1 Graphique 1 (p.8)

```
revenus_a_imputer <- bdd %>%
  mutate(sdrevcl = ifelse(sdrevcl==999999999,NA,sdrevcl)) %>%
  group_by(sdrevtr) %>%
  summarise(sdrevcl = mean(sdrevcl,na.rm=TRUE))

bdd <- bdd %>%
  mutate(
    sdrevcl_imput_moy = case_when(
```

```

sdrevcl!=999999999 ~ sdrevcl,
sdrevcl==999999999 & sdrevtr==1 ~ revenus_a_imputer$sdrevcl[1],
sdrevcl==999999999 & sdrevtr==2 ~ revenus_a_imputer$sdrevcl[2],
sdrevcl==999999999 & sdrevtr==3 ~ revenus_a_imputer$sdrevcl[3],
sdrevcl==999999999 & sdrevtr==4 ~ revenus_a_imputer$sdrevcl[4],
sdrevcl==999999999 & sdrevtr==5 ~ revenus_a_imputer$sdrevcl[5],
sdrevcl==999999999 & sdrevtr==6 ~ revenus_a_imputer$sdrevcl[6],
sdrevcl==999999999 & sdrevtr==7 ~ revenus_a_imputer$sdrevcl[7]
)
)

bdd <- bdd %>%
  mutate(seuil_pauvrete = do.call(c,lapply(annee,function(annee_selec){
0.6*median(bdd[bdd$annee==annee_selec,"sdrevcl_imput_moy"]/bdd[bdd$annee==annee_selec,"sduc"],
na.rm=TRUE)
})
))) %>%
  mutate(subj_inf_mini_decla = sdrevcl_imput_moy - ifelse(pe16==999999999,NA,pe16) <0) %>%
  mutate(subj_pauvrete =
    ifelse(pe3==3, TRUE,ifelse(pe3==4,NA,FALSE))) %>%
  mutate(subj_besoin_aide_etat = ifelse(pe15==2, TRUE,FALSE)) %>%
  mutate(obj_pauvrete= sdrevcl_imput_moy/sduc<seuil_pauvrete)

data_g1 <- bdd %>%
  group_by(annee) %>%
  summarise(subj_pauvrete = round(100*weighted.mean(subj_pauvrete,
poids, na.rm=TRUE)),
    subj_inf_mini_decla = round(100*weighted.mean(subj_inf_mini_decla,
poids, na.rm=TRUE)),
    obj_pauvrete = round(100*weighted.mean(obj_pauvrete,
poids, na.rm=TRUE)),
    subj_besoin_aide_etat = round(100*weighted.mean(subj_besoin_aide_etat,
poids, na.rm=TRUE))
  )

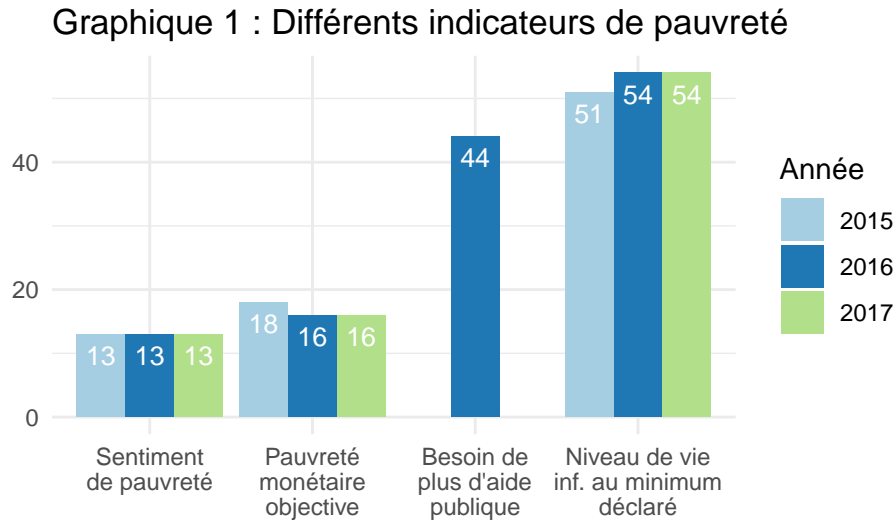
ggplot(data=data_g1 %>% tidyr::pivot_longer(!annee),
  aes(x=name, y=value, fill=annee)) +
  geom_bar(stat="identity", position=position_dodge())+
  geom_text(aes(label=value), vjust=1.6, color="white",
    position = position_dodge(0.9), size=3.5)+
  scale_fill_brewer(palette="Paired")+
  theme_minimal() +
  scale_x_discrete(name="",
    limits=c("subj_pauvrete","obj_pauvrete",
"subj_besoin_aide_etat",
"subj_inf_mini_decla"),
    labels=c(

```

```

"obj_pauvrete" = "Pauvreté\nmonétaire\nobjective",
"subj_inf_mini_decla" = "Niveau de vie\ninf. au minimum\ndéclaré",
"subj_besoin_aide_etat"="Besoin de\nplus d'aide\npublique",
"subj_pauvrete" = "Sentiment\nde pauvreté"))+
ggtitle("Graphique 1 : Différents indicateurs de pauvreté") + ylab("")+
labs(fill="Année")

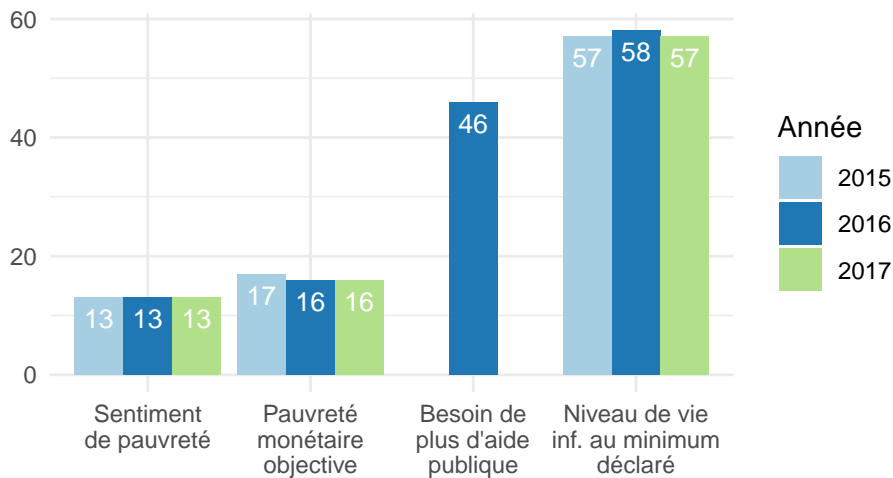
```



Commentaires :

- Les NSP sont exclus du champ pour `subj_pauvrete`, pour `subj_inf_mini_decla` mais apparemment pas pour `subj_besoin_aide_etat`. Il faudrait le faire aussi (correction dans le graphique qui suit).
- Il est intéressant de voir que le fait de prendre en compte ou non le “0” dans la différence entre le revenu réel et le minimum pour vivre fait varier significativement l’indicateur `subj_inf_mini_decla` (l’augmenterait d’une dizaine de points)
- Il est possible d’imputer le revenu d’une autre façon que celle proposée dans l’article de Duvoux et Papuchon (ils proposent l’imputation par la moyenne des revenus en clairs par tranche). Nous avons utilisé une autre méthode d’imputation (documentation à préciser) pour tester dans le graphique ci-après.

Graphique 1 (corrigé) : Différents indicateurs de pauvreté



### 3.2 Graphique 2 (p.16) et Tableau 2 (p.17)

```
bdd <- bdd %>%
  mutate(
    prof_exercee_ou_derniere = case_when(
      sdpcs7%in%c(1,2) ~ "Agriculteurs, artisans et commerçants",
      sdpcs7%in%c(3) ~ "Cadre supérieur, profession libérale",
      sdpcs7%in%c(4) ~ "Profession intermédiaire",
      sdpcs7%in%c(5) ~ "Employé",
      sdpcs7%in%c(6) ~ "Ouvrier",
      sdpcs7%in%c(7) ~ "Autres (n'ayant jamais travaillé)"
    )
  ) %>%
  mutate(
    activite_actuelle = case_when(
      sdsitua==4 ~ "Chômeur",
      sdsitua==6 ~ "Retraité",
      (sdsitua==1 & sdstatemp!=1) | sdsitua%in%c(2,3) ~
        "Emploi précaire/temps partiel",
      TRUE #sdsitua%in%c(5,7) | (sdsitua==1 & sdstatemp==1) incomplet
      ~
        "Autres (CDI à temps plein, étudiants en formations et personnes n'ayant jamais travaillé)"
    )
  ) %>%
  mutate(
    PCS_recode = case_when(
      prof_exercee_ou_derniere%in%c("Employé","Ouvrier") & activite_actuelle=="Emploi précaire/temps partiel" ~
        "Employés et ouvriers\nprécaires ou\ntemps partiels",
      prof_exercee_ou_derniere%in%c("Employé","Ouvrier") & activite_actuelle=="Retraité" ~
        "Employés et ouvriers\nà la retraite", #sdsitua==6
      prof_exercee_ou_derniere%in%c("Employé","Ouvrier") & activite_actuelle=="Chômeur" ~
```

```

      "Employés et ouvriers\nau chômage",
      sdsitua==1 & !(prof_exercee_ou_derniere%in%c("Employé","Ouvrier") & activite_actuelle=="En
      sdsitua%in%c(7) ~ "Sans activité\nprofessionnelle",
      TRUE ~ "Autres"
    )
  )
)

```

```

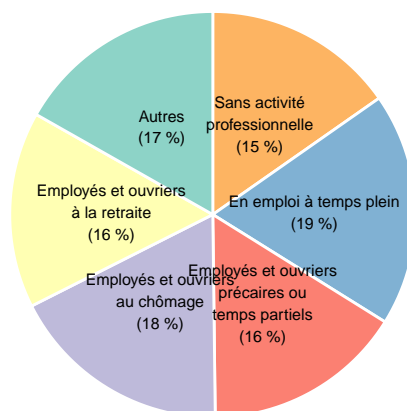
tab_croise_g2 <- questionr::cprop(questionr::wtd.table(
  bdd$PCS_recode,bdd$subj_pauvrete, weights = bdd$poids))

data_g2 <- data.frame(
  group=row.names(tab_croise_g2)[-nrow(tab_croise_g2)],
  value=tab_croise_g2[,2][-nrow(tab_croise_g2)]
) %>%
  arrange(desc(group)) %>%
  mutate(ypos = cumsum(value)- 0.5*value) %>%
  mutate(labels=paste0(group,"\n(",round(value,0)," %)")

ggplot(data_g2, aes(x="", y=value, fill=group)) +
  geom_bar(stat="identity", width=1, color="white") +
  coord_polar("y", start=0) +
  theme_void() +
  theme(legend.position="none",
        plot.title = element_text(size=10)) +
  geom_text(aes(y = ypos, label = labels), color = "black", size=2) +
  scale_fill_brewer(palette="Set3") +
  labs(title = "Graphique 2 : Qui sont les personnes\nqui se sentent pauvres ?")

```

Graphique 2 : Qui sont les personnes  
qui se sentent pauvres ?



```

tab_croise_t2a <- questionr::cprop(questionr::wtd.table(bdd$prof_exercee_ou_derniere,
  bdd$subj_pauvrete,
  weights = bdd$poids)
  )[c(1,3,6,4,5,2,7),c(2,1)]

```

TABLE 1 – Tableau 2 - Statut d'activité et dernière profession exercée des personnes qui se disent pauvres, par rapport au reste de la population

	Se sent pauvre	Ne se sent pas pauvre
Agriculteurs, artisans et commerçants	6	7
Cadre supérieur, profession libérale	4	17
Profession intermédiaire	9	21
Employé	33	25
Ouvrier	28	19
Autres (n'ayant jamais travaillé)	20	11
Total	100	100
Chômeur	22	8
Retraité	22	31
Emploi précaire/temps partiel	19	13
Autres (CDI à temps plein, étudiants en formations et personnes n'ayant jamais travaillé)	37	48
Total	100	100

```
tab_croise_t2b <- questionr::cprop(questionr::wtd.table(bdd$activite_actuelle,
                                                         bdd$subj_pauvrete,
                                                         weights = bdd$poids)
                                ) [c(2,4,3,1,5),c(2,1)]

data_t2 <- round(rbind(tab_croise_t2a,tab_croise_t2b),0)
colnames(data_t2)<-c("Se sent pauvre","Ne se sent pas pauvre")

data_t2 %>%
  kable(caption="Tableau 2 - Statut d'activité et dernière profession exercée des personnes qui
  kableExtra::column_spec(1, width = "9cm") %>%
  kableExtra::column_spec(2, width = "3cm") %>%
  kableExtra::column_spec(3, width = "3cm")
```

— Il y a des petits écarts d'un point ou deux pour le camembert. Le codage n'est pas très clair (ce qui est compté dans autres et dans sans activité professionnelle).

### 3.3 Graphiques 3 (p.23), 4 (p.24) et 5 (p.25)

```
bdd <- bdd %>%
  mutate(
    sitfam = case_when(
      sdsitfam==2 ~ "Membre du couple",
      sdsitfam==1 ~ "Vie seule",
      sdsitfam==3 ~ "Famille monoparentale",
      sdsitfam%in%c(4,5,6,7) ~ "Autre situation (enfant de la famille,\nami ou parent hébergé, c
      #pas précisé si prise compte de 7 mais très peu d'individus
    )
```

```

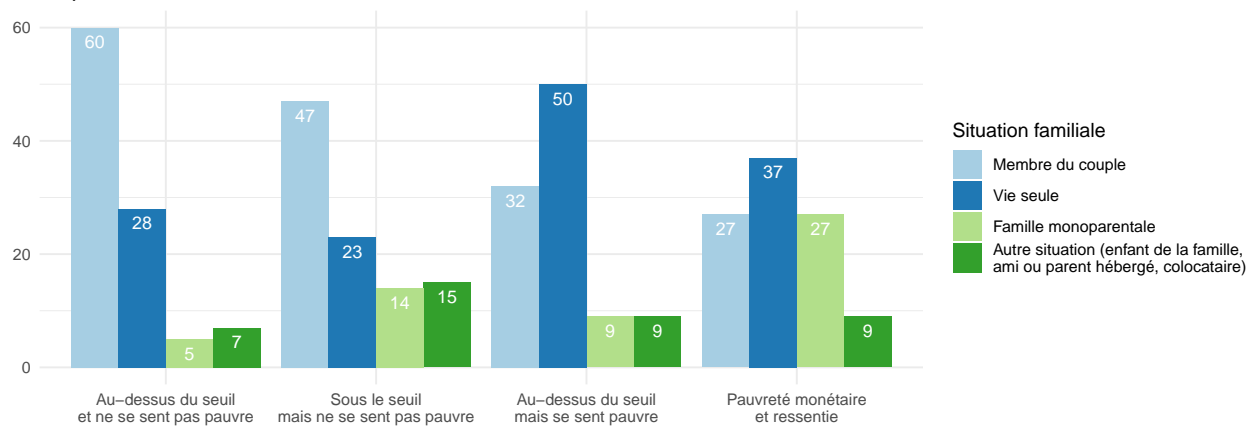
) %>%
mutate(
  crois_pauvrete_obj_subj = case_when(
    !subj_pauvrete & !obj_pauvrete & !is.na(subj_pauvrete) & !is.na(obj_pauvrete) ~
      "Au-dessus du seuil\net ne se sent pas pauvre",
    !subj_pauvrete & obj_pauvrete & !is.na(subj_pauvrete) & !is.na(obj_pauvrete) ~
      "Sous le seuil\nmais ne se sent pas pauvre",
    subj_pauvrete & !obj_pauvrete & !is.na(subj_pauvrete) & !is.na(obj_pauvrete) ~
      "Au-dessus du seuil\nmais se sent pauvre",
    subj_pauvrete & obj_pauvrete & !is.na(subj_pauvrete) & !is.na(obj_pauvrete) ~
      "Pauvreté monétaire\net ressentie"
  )
)

data_g3 <- data.frame(
  round(
    questionr::cprop(
      questionr::wtd.table(bdd$sitfam,
                           bdd$crois_pauvrete_obj_subj,
                           weights=bdd$poids)
    )[-5,-5]
  )
) %>% mutate(Var1 = factor(Var1, levels=c(
  "Membre du couple", "Vie seule",
  "Famille monoparentale",
  "Autre situation (enfant de la famille,\nami ou parent hébergé, colocataire)"
)))

ggplot(data=data_g3,
  aes(x=Var2, y=Freq, fill=Var1)) +
  geom_bar(stat="identity", position=position_dodge())+
  geom_text(aes(label=Freq), vjust=1.6, color="white",
    position = position_dodge(0.9), size=3.5)+
  scale_fill_brewer(palette="Paired")+
  theme_minimal() +
  scale_x_discrete(name="",
    limits=c("Au-dessus du seuil\net ne se sent pas pauvre",
             "Sous le seuil\nmais ne se sent pas pauvre",
             "Au-dessus du seuil\nmais se sent pauvre",
             "Pauvreté monétaire\net ressentie"))+
  ggtitle("Graphique 3 : Situation familiale en fonction des indicateurs
de pauvreté monétaire et ressentie") + ylab("")+
  labs(fill="Situation familiale")

```

Graphique 3 : Situation familiale en fonction des indicateurs de pauvreté monétaire et ressentie

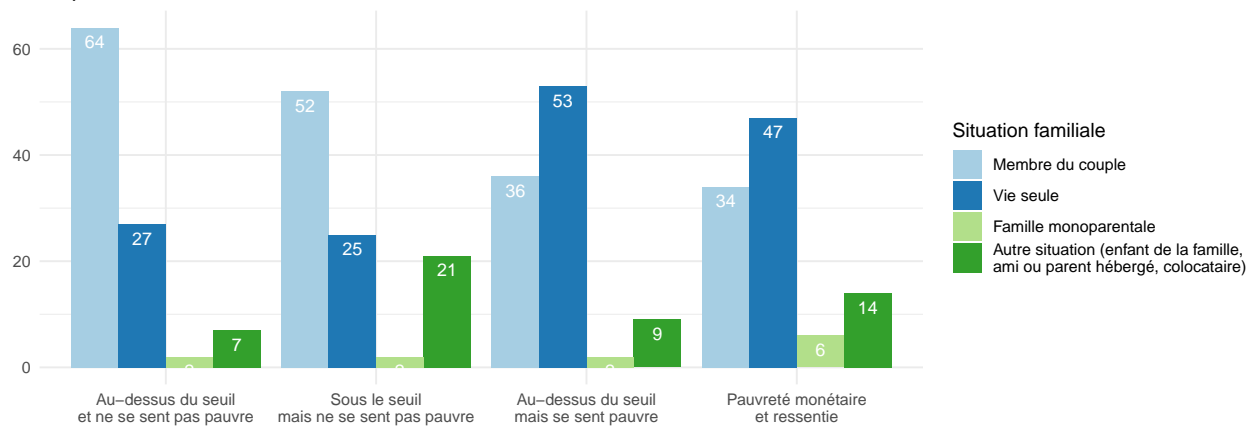


```
data_g4 <- data.frame(
  round(
    questionr::cprop(
      questionr::wtd.table(bdd$sitfam[bdd$sdsexe==1],
                           bdd$crois_pauvrete_obj_subj[bdd$sdsexe==1],
                           weights=bdd$poids[bdd$sdsexe==1])
    )[-5,-5]
  )
) %>% mutate(Var1 = factor(Var1, levels=c(
  "Membre du couple", "Vie seule",
  "Famille monoparentale",
  "Autre situation (enfant de la famille,\nami ou parent hébergé, colocataire)"
)))

ggplot(data=data_g4,
       aes(x=Var2, y=Freq, fill=Var1)) +
  geom_bar(stat="identity", position=position_dodge()) +
  geom_text(aes(label=Freq), vjust=1.6, color="white",
            position = position_dodge(0.9), size=3.5) +
  scale_fill_brewer(palette="Paired") +
  theme_minimal() +
  scale_x_discrete(name="",
                   limits=c("Au-dessus du seuil\net ne se sent pas pauvre",
                             "Sous le seuil\nmais ne se sent pas pauvre",
                             "Au-dessus du seuil\nmais se sent pauvre",
                             "Pauvreté monétaire\net ressentie")) +
  ggtitle("Graphique 4 : Situation familiale DES HOMMES en fonction des indicateurs
de pauvreté monétaire et ressentie") + ylab("") +
  labs(fill="Situation familiale")
```



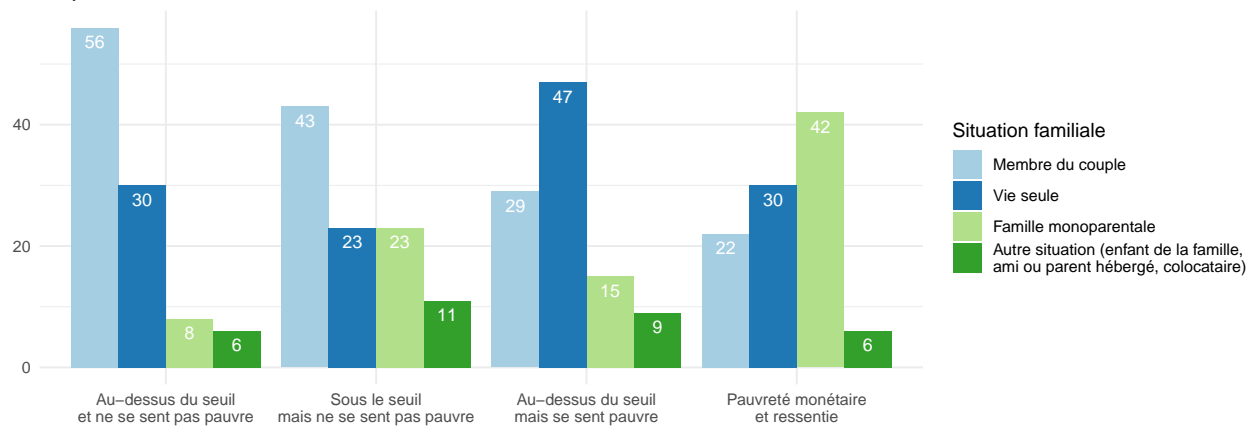
Graphique 4 : Situation familiale DES HOMMES en fonction des indicateurs de pauvreté monétaire et ressentie



```
data_g4 <- data.frame(
  round(
    questionr::cprop(
      questionr::wtd.table(bdd$sitfam[bdd$sdsexe==2],
                           bdd$crois_pauvrete_obj_subj[bdd$sdsexe==2],
                           weights=bdd$poids[bdd$sdsexe==2])
    )[-5,-5]
  )
) %>% mutate(Var1 = factor(Var1, levels=c(
  "Membre du couple", "Vie seule",
  "Famille monoparentale",
  "Autre situation (enfant de la famille,\nami ou parent hébergé, colocataire)"
)))

ggplot(data=data_g4,
       aes(x=Var2, y=Freq, fill=Var1)) +
  geom_bar(stat="identity", position=position_dodge()) +
  geom_text(aes(label=Freq), vjust=1.6, color="white",
            position = position_dodge(0.9), size=3.5) +
  scale_fill_brewer(palette="Paired") +
  theme_minimal() +
  scale_x_discrete(name="",
                   limits=c("Au-dessus du seuil\net ne se sent pas pauvre",
                             "Sous le seuil\nmais ne se sent pas pauvre",
                             "Au-dessus du seuil\nmais se sent pauvre",
                             "Pauvreté monétaire\net ressentie")) +
  ggtitle("Graphique 5 : Situation familiale DES FEMMES en fonction des indicateurs
de pauvreté monétaire et ressentie") + ylab("") +
  labs(fill="Situation familiale")
```

Graphique 5 : Situation familiale DES FEMMES en fonction des indicateurs de pauvreté monétaire et ressentie



— Encore des écarts d'un point maximum par rapport aux données de l'article.

### 3.4 Tableaux 4 (p.26) et 5 (p.27)

— Dans l'article, la notion d'assistance est décrite comme la réception du RSA, d'allocation chômage (mais cet élément est absent des régressions, seul le statut chômeur est mentionné) ou de prestations liées au handicap, validité et dépendance. L'aide au logement n'est pas prise en compte

```
#sdres_3 : RSA, sdres_4 : chômage, sdres_10 : handicap dépdce.
```

```
bdd <- bdd %>%
  mutate(
    assistance = case_when(
      sdres_3==1 | sdres_4==1 | sdres_10==1 ~ TRUE,
      sdres_3==2 & sdres_4==2 & sdres_10==2 ~ FALSE,
      TRUE ~ NA
    )
  ) %>%
  mutate(
    pessimisme = case_when(
      og3_1%in%c(1,2) ~ FALSE,
      og3_1%in%c(3,4) ~ TRUE,
      og3_1%in%c(5) ~ NA
    )
  ) %>%
  mutate(
    declassement = case_when(
      og2_ab%in%c(1,2) | og2_cd%in%c(4,5) ~ TRUE,
      og2_ab%in%c(3,4,5) | og2_cd%in%c(1,2,3) ~ FALSE,
      og2_ab%in%c(6) | og2_cd%in%c(6) ~ NA
    )
  )
```

```
data_t4 <- data.frame(round(cbind(
  #colonne1
```

TABLE 2 – Tableau 4 : Pauvreté et trajectoire sociale escomptée

	Sous le seuil de pauvreté	Au-dessus du seuil de pauvreté	Pas en situation d'assistance	Situation d'assistance	Ensemble
Ne se déclare pas pauvre	45	39	38	44	40
Se déclare pauvre	62	63	60	65	63
Ensemble	51	41	40	49	43

```

questionr::rprop(questionr::wtd.table(bdd$subj_pauvrete[bdd$obj_pauvrete],
                                      bdd$pessimisme[bdd$obj_pauvrete],
                                      weights=bdd$poids[bdd$obj_pauvrete]))[c(1,2,3),2],
#colonne2
questionr::rprop(questionr::wtd.table(bdd$subj_pauvrete[!bdd$obj_pauvrete],
                                      bdd$pessimisme[!bdd$obj_pauvrete],
                                      weights=bdd$poids[!bdd$obj_pauvrete]))[c(1,2,3),2],
#colonne3
questionr::rprop(questionr::wtd.table(bdd$subj_pauvrete[!bdd$assistance],
                                      bdd$pessimisme[!bdd$assistance],
                                      weights=bdd$poids[!bdd$assistance]))[c(1,2,3),2],
#colonne4
questionr::rprop(questionr::wtd.table(bdd$subj_pauvrete[bdd$assistance],
                                      bdd$pessimisme[bdd$assistance],
                                      weights=bdd$poids[bdd$assistance]))[c(1,2,3),2],
#colonne5
questionr::rprop(questionr::wtd.table(bdd$subj_pauvrete,
                                      bdd$pessimisme,
                                      weights=bdd$poids))[c(1,2,3),2]
)))
row.names(data_t4) <- c("Ne se déclare pas pauvre", "Se déclare pauvre", "Ensemble")
colnames(data_t4) <- c("Sous le seuil de pauvreté", "Au-dessus du seuil de pauvreté",
                      "Pas en situation d'assistance", "Situation d'assistance", "Ensemble")
data_t4 %>%
  kable(caption="Tableau 4 : Pauvreté et trajectoire sociale escomptée") %>%
  kableExtra::column_spec(1, width = "4cm") %>%
  kableExtra::column_spec(2:5, width = "2cm")

data_t5 <- data.frame(round(cbind(
#colonne1
questionr::rprop(questionr::wtd.table(bdd$subj_pauvrete[bdd$obj_pauvrete],
                                      bdd$declasserment[bdd$obj_pauvrete],
                                      weights=bdd$poids[bdd$obj_pauvrete]))[c(1,2,3),2],
#colonne2
questionr::rprop(questionr::wtd.table(bdd$subj_pauvrete[!bdd$obj_pauvrete],
                                      bdd$declasserment[!bdd$obj_pauvrete],
                                      weights=bdd$poids[!bdd$obj_pauvrete]))[c(1,2,3),2],

```

```

#colonne3
questionr::rprop(questionr::wtd.table(bdd$subj_pauvrete[!bdd$assistance],
                                      bdd$declassement[!bdd$assistance],
                                      weights=bdd$poids[!bdd$assistance]))[c(1,2,3),2],

#colonne4
questionr::rprop(questionr::wtd.table(bdd$subj_pauvrete[bdd$assistance],
                                      bdd$declassement[bdd$assistance],
                                      weights=bdd$poids[bdd$assistance]))[c(1,2,3),2],

#colonne5
questionr::rprop(questionr::wtd.table(bdd$subj_pauvrete,
                                      bdd$declassement,
                                      weights=bdd$poids))[c(1,2,3),2]
)))
row.names(data_t5) <- c("Ne se déclare pas pauvre", "Se déclare pauvre", "Ensemble")
colnames(data_t5) <- c("Sous le seuil de pauvreté", "Au-dessus du seuil de pauvreté",
                      "Pas en situation d'assistance", "Situation d'assistance", "Ensemble")
data_t5 %>%
  kable(caption="Tableau 5 : Pauvreté et trajectoire sociale perçue")%>%
  kableExtra::column_spec(1, width = "4cm") %>%
  kableExtra::column_spec(2:5, width = "2cm")

```

— Encore des écarts par rapport aux données de l'article (monte jusqu'à deux points).

### 3.5 Tableau 6 (p.28)

```

bdd <- bdd %>%
  mutate(
    proprietaire = case_when(
      lo1==1 ~ TRUE,
      lo1%in%c(2,3,4) ~ FALSE,
      lo1==5 ~ NA
    )
  ) %>%
  mutate(
    retraite = case_when(
      sdsitua==6 ~ TRUE,
      TRUE ~ FALSE
    )
  )

data_t6 <- data.frame(round(cbind(
#colonne1
questionr::rprop(questionr::wtd.table(bdd$proprietaire[bdd$retraite],
                                      bdd$subj_pauvrete[bdd$retraite],
                                      weights=bdd$poids[bdd$retraite]))[c(2,1,3),2],

#colonne2
questionr::rprop(questionr::wtd.table(bdd$proprietaire[!bdd$retraite],
                                      bdd$subj_pauvrete[!bdd$retraite],
                                      weights=bdd$poids[!bdd$retraite]))[c(2,1,3),2]

```

TABLE 3 – Tableau 6 : Les retraités non propriétaires particulièrement surexposés au sentiment de pauvreté

	Retraités	Non-retraités
Propriétaires	3	5
Non-propriétaires	21	20
Ensemble	9	14

```
)))
row.names(data_t6) <- c("Propriétaires", "Non-propriétaires", "Ensemble")

colnames(data_t6) <- c("Retraités", "Non-retraités")

data_t6 %>%
  kable(caption="Tableau 6 : Les retraités non propriétaires particulièrement surexposés au sent
  kableExtra::column_spec(1, width = "4cm") %>%
  kableExtra::column_spec(2:3, width = "3cm")
```

— Mêmes pourcentages que dans l'article

### 3.6 Tableau 1 – Modèle 1 (p.13) et Tableau 3 – Modèle 2 (p.22)

```
bdd_logit <- bdd %>%
  mutate(quantile_nivie = factor(
    cut(sdrevcl_imput_moy/sduc, include.lowest=TRUE,
      breaks=quantile(sdrevcl_imput_moy/sduc,
        probs=seq(0,1,0.2),na.rm=TRUE),
        labels=1:5),
    levels=c(2,1,3,4,5),
    labels=paste0("-Quintile ",c(2,1,3,4,5)))
  ) %>%
  mutate(
    statact = factor(case_when(
      sdsitua==1 & sdstatemp%in%c(1) ~ 1,
      sdsitua%in%c(2,3) | (sdsitua==1 & sdstatemp%in%c(NA,2,3,4,5)) ~ 2,
      sdsitua==4 ~ 3,
      sdsitua==5 ~ 4,
      sdsitua==6 ~ 5,
      sdsitua==7 ~ 6,
      TRUE ~ NA_real_
    ),
    levels=1:6,
    labels=paste0("-",
      c("CDI temps plein",
        "Emploi précaire ou à temps partiel",
        "Recherche d'emploi",
```

```

        "Étudiant",
        "Retraité",
        "Aucune activité professionnelle"))
  )) %>%
mutate(prof = factor(
  #ifelse(sdpcs7!=7,sdpcs7,NA),
  sdpcs7,
  levels=c(4,1,2,3,5,6,7),
  labels=paste0("-",c("Profession intermédiaire","Agriculteur",
    "Artisan commerçant","Cadre supérieur, profession libérale",
    "Employé","Ouvrier","Autre inactif")))
)) %>%
# fastDummies::dummy_cols(select_columns = 'prof') %>%
mutate(
  diplome = factor(case_when(
    sddipl%in%c(1,2,3,4) ~ 1,
    sddipl%in%c(5,6) ~ 2,
    sddipl%in%c(7) ~ 3,
    sddipl%in%c(8) ~ 4,
    TRUE ~ NA_real_
  ),
  levels=c(2,1,3,4),
  labels=paste0("-",c("Baccalauréat","CAP, BEP ou moins","Bac + 2","Bac + 3 ou plus")))
)) %>%
mutate(aide_log = factor(
  ifelse(sdres_9!=3,sdres_9,NA),
  levels=c(2,1),
  labels=paste0("-",c("Pas d'aide au logement","Aide au logement reçue")))
)) %>%
mutate(aide_rsa = factor(
  ifelse(sdres_3!=3,sdres_3,NA),
  levels=c(2,1),
  labels=paste0("-",c("Pas de RSA","RSA reçu")))
)) %>%
mutate(aide_handi = factor(
  ifelse(sdres_10!=3,sdres_10,NA),
  levels=c(2,1),
  labels=paste0("-",c("Pas d'alloc. hand./invalid./dépend.",
    "Alloc. hand./invalid./dépend. reçu")))
)) %>%
mutate(sexe = factor(
  ifelse(sdsexe!=3,sdsexe,NA),
  levels=c(2,1),
  labels=paste0("-",c("Femme","Homme")))
)) %>%
mutate(age_tranche = factor(
  cut(sdage,
    breaks=c(18,30,40,50,60,70,120),

```

```

      include.lowest=TRUE,
      labels=1:6),
    levels=c(2,1,3,4,5,6),
    labels=paste0("-",c("30 à 39 ans","18 à 29 ans","40 à 49 ans",
      "50 à 59 ans","60 à 69 ans", "70 ans et plus"))))
) %>%
mutate(
  vie_fam = factor(case_when(
    sdsitfam==1 ~ 1,
    sdsitfam==2 & sdnbenf==0 ~ 2,
    sdsitfam==2 & sdnbenf!=0 ~ 3,
    sdsitfam==3 ~ 4,
    sdsitfam==4 ~ 5,
    sdsitfam%in%c(5,6,7) ~ 6,
    TRUE ~ NA_real_
  )),
  levels=c(2,1,3,4,5,6),
  labels=paste0("-",
    c("Membre du couple (pas d'enfants à charge)","Vit seul",
      "Membre du couple (enfants à charge)", "Chef famille monoparentale",
      "Enfant","Autre")))
)) %>%
mutate(
  statut_occup = factor(case_when(
    lo1%in%c(2,3,4) ~ 1,
    lo1==1 ~ 2,
    lo1==5 ~ NA_real_
  )),
  levels=c(1,2),
  labels=paste0("-",
    c("Locataire ou hébergé","Propriétaire")))
)) %>%
mutate(annee_fac = factor(annee,labels=paste0("-",2015:2017))) %>%
mutate(pauvrete_mon_rel= factor(
  ifelse(sdrevcl_imput_moy/sduc<seuil_pauvrete,3,
    ifelse(sdrevcl_imput_moy/sduc>=1.2*seuil_pauvrete,1,
      ifelse(sdrevcl_imput_moy/sduc<1.2*seuil_pauvrete &
        sdrevcl_imput_moy/sduc>=seuil_pauvrete ,
        2,NA)
    )),
  levels=c(1,2,3),
  labels=paste0("-",c("Revenus > 20 pourcents seuil pauvreté",
    "Entre le seuil et 20 pourcents au dessus",
    "Pauvreté monétaire")))
)
)

```

Modèle 1

```

reg1 <- glm(subj_pauvrete ~ quantile_nivie + statact +
  prof +
  diplome + aide_log +
  aide_rsa + aide_handi + sexe + age_tranche + vie_fam +
  statut_occup + annee_fac,
  data = bdd_logit %>% select(`subj_pauvrete`, `quantile_nivie`,
    `statact`, `prof`, `diplome`, `aide_log`, `aide_rsa`,
    `aide_handi`, `sexe`, `age_tranche`,
    `vie_fam`, `statut_occup`, `annee_fac`
  ) %>% tidyr::drop_na(),
  family = binomial(logit))
results1 <- summary(reg1)$coefficients
modalites1 <- c("(Intercept)")
for(var in c("quantile_nivie", "statact", "prof",
  "diplome", "aide_log", "aide_rsa",
  "aide_handi", "sexe", "age_tranche",
  "vie_fam", "statut_occup", "annee_fac")){
  modalites1 <- c(modalites1, gsub("-", "", levels(bdd_logit[,var])))
}

df_model1 <- data.frame(noms= c("-(Intercept)", row.names(results1)[-1]),
  valeur=round(results1[,1],2),
  pvaleur=cut(results1[,4], breaks = c(0, 0.001, 0.01, 0.05, 1),
    include.lowest = T,
    labels = c('***', '**', '*', '')),
  row.names=NULL
) %>% mutate(noms = gsub("-", "", noms)) %>%
  tidyr::separate(noms, c("variable", "modalite"), sep="-") %>%
  mutate(odds=ifelse(is.na(pvaleur) | pvaleur=="", NA, round(exp(results1[,1]),2))) %>%
  right_join(data.frame(modalite=modalites1), by="modalite") %>%
  arrange(match(modalite, modalites1))

R2_ajuste1 <- with(summary(reg1), 1 - deviance/null.deviance)
N1 <- length(summary(reg1)$deviance.resid)

df_model1 %>% select(-variable) %>%
  kable(format = 'latex', booktabs = TRUE, longtable = TRUE,
  caption="Tableau 1 - Les déterminants du sentiment de pauvreté (Modèle 1)
") %>%
  kableExtra::column_spec(1, width = "4cm") %>%
  kableExtra::column_spec(2:4, width = "2cm") %>%
  kableExtra::kable_styling(latex_options = c("hold_position", "repeat_header")) %>%
  kableExtra::footnote(paste0("Modèle logit (Variable dépendante = Se déclarer pauvre) / N =
    N1, / R2 ajusté = ",
    round(100*R2_ajuste1,1), "%"))

```



TABLE 4: Tableau 1 - Les déterminants du sentiment de  
pauvreté (Modèle 1)

modalite	valeur	pvalueur	odds
(Intercept)	-2.38	***	0.09
Quintile 2			
Quintile 1	0.50	***	1.65
Quintile 3	-0.73	***	0.48
Quintile 4	-1.44	***	0.24
Quintile 5	-1.74	***	0.18
CDI temps plein			
Emploi précaire ou à temps partiel	0.34	**	1.41
Recherche d'emploi	0.43	**	1.54
Étudiant	0.47		
Retraité	-0.09		
Aucune activité professionnelle	1.23	***	3.42
Profession intermédiaire			
Agriculteur	0.32		
Artisan commerçant	0.23		
Cadre supérieur, profession libérale	-0.40		
Employé	0.58	***	1.78
Ouvrier	0.29		
Autre inactif	-0.34		
Baccalauréat			
CAP, BEP ou moins	0.05		
Bac + 2	-0.20		
Bac + 3 ou plus	-0.19		
Pas d'aide au logement			
Aide au logement reçue	0.32	***	1.37
Pas de RSA			
RSA reçu	0.63	***	1.87
Pas d'alloc. hand./invalid./dépend.			
Alloc. hand./invalid./dépend.	0.10		
reçu			
Femme			
Homme	0.30	***	1.35
30 à 39 ans			
18 à 29 ans	-0.40	**	0.67
40 à 49 ans	0.24	*	1.27
50 à 59 ans	0.11		

TABLE 4: Tableau 1 - Les déterminants du sentiment de pauvreté (Modèle *(continued)*)

modalite	valeur	pvalue	odds
60 à 69 ans	0.29		
70 ans et plus	0.23		
Membre du couple (pas d'enfants à charge)			
Vit seul	0.70	***	2.01
Membre du couple (enfants à charge)	-0.34	*	0.71
Chef famille monoparentale	0.47	**	1.59
Enfant	0.29		
Autre	0.73	**	2.07
Locataire ou hébergé			
Propriétaire	-0.80	***	0.45
2015			
2016	-0.09		
2017	-0.08		

*Note:*

Modèle logit (Variable dépendante = Se déclarer pauvre) / N = 8460 / R2 ajusté = 24.8 %

## Modèle 2

```
#On remplace la tranche d'âge par l'âge
#On remplace les quantiles de niveau de vie par pauvreté monétaire relative
#On enlève l'assistance
reg2 <- glm(subj_pauvrete ~ sdage + pauvreté_mon_rel + diplome + statact +
             sexe + prof + vie_fam + statut_occup + annee_fac,
             data = bdd_logit %>% select(`subj_pauvrete`, `sdage`, `pauvrete_mon_rel`, `statact`, `p
                                   `vie_fam`, `statut_occup`, `annee_fac`
             ) %>% tidyr::drop_na(),
             family = binomial(logit))

results2 <- summary(reg2)$coefficients
modalites2 <- c("(Intercept)", "sdage")
for(var in c("sdage", "pauvrete_mon_rel", "diplome", "statact", "sexe",
             "prof", "vie_fam", "statut_occup", "annee_fac")){
  modalites2 <- c(modalites2, gsub("-", "", levels(bdd_logit[,var])))
}

df_model2 <- data.frame(noms= c("-(Intercept)", "-sdage", row.names(results2)[-c(1,2)]),
                       valeur=round(results2[,1],2),
                       pvalue=cut(results2[,4], breaks = c(0, 0.001, 0.01, 0.05, 1),
                                include.lowest = T,
```

```

                                labels = c('***', '**', '*', '')),
                                row.names=NULL
) %>% mutate(noms = gsub("`", "", noms)) %>%
tidyr::separate(noms, c("variable", "modalite"), sep="-") %>%
mutate(odds=ifelse(is.na(pvaleur) | pvaleur=="", NA, round(exp(results2[,1]), 2))) %>%
right_join(data.frame(modalite=modalites2), by="modalite") %>%
arrange(match(modalite, modalites2))

R2_ajuste2 <- with(summary(reg2), 1 - deviance/null.deviance)
N2 <- length(summary(reg2)$deviance.resid)

df_model2 %>% select(-variable) %>%
kable(format = 'latex', booktabs = TRUE, longtable = TRUE,
caption="Tableau 4 - Facteurs du sentiment de pauvreté (Modèle 2)") %>%
kableExtra::column_spec(1, width = "4cm") %>%
kableExtra::column_spec(2:4, width = "2cm") %>%
kableExtra::kable_styling(latex_options = c("hold_position", "repeat_header")) %>%
kableExtra::footnote(paste0("Modèle logit (Variable dépendante = Se déclarer pauvre) / N =
                                N2," / R2 ajusté = ",
                                round(100*R2_ajuste2, 1), " %"))

```

TABLE 5: Tableau 4 - Facteurs du sentiment de pauvreté  
(Modèle 2)

modalite	valeur	pvaleur	odds
(Intercept)	-3.85	***	0.02
sdage	0.01	**	1.01
Revenus > 20 pourcents seuil pauvreté			
Entre le seuil et 20 pourcents au dessus	0.91	***	2.48
Pauvreté monétaire	1.41	***	4.10
Baccalauréat			
CAP, BEP ou moins	0.10		
Bac + 2	-0.24		
Bac + 3 ou plus	-0.30		
CDI temps plein			
Emploi précaire ou à temps partiel	0.57	***	1.76
Recherche d'emploi	0.71	***	2.03
Étudiant	0.55		
Retraité	-0.07		
Aucune activité professionnelle	1.56	***	4.75
Femme			
Homme	0.30	***	1.35

TABLE 5: Tableau 4 - Facteurs du sentiment de pauvreté  
(Modèle 2) (*continued*)

modalite	valeur	pvalue	odds
Profession intermédiaire			
Agriculteur	0.56		
Artisan commerçant	0.22		
Cadre supérieur, profession libérale	-0.53	*	0.59
Employé	0.70	***	2.01
Ouvrier	0.46	**	1.59
Autre inactif	-0.23		
Membre du couple (pas d'enfants à charge)			
Vit seul	0.91	***	2.50
Membre du couple (enfants à charge)	-0.07		
Chef famille monoparentale	0.90	***	2.47
Enfant	0.29		
Autre	0.72	**	2.06
Locataire ou hébergé			
Propriétaire	-1.06	***	0.35
2015			
2016	-0.06		
2017	-0.06		

*Note:*

Modèle logit (Variable dépendante = Se déclarer pauvre) / N = 8468 / R2 ajusté = 22.5 %

- Les résultats principaux sont les mêmes mais les coefficients parfois légèrement différents.
- Le modèle tel qu'exposé dans Papuchon et Duvoux n'est pas estimable dans le sens où les personnes sans PCS sont censées être exclues du champ. Or les personnes sans activité professionnelle le sont bien dans la partie statut d'activité. On ajoute donc une catégorie autres dans les PCS.
- Dans quel case ranger un indépendant ? CDI à temps plein ? Précaire à temps partiel ? Car les indépendants ne répondent pas sur leur contrat de travail.

## 4 Notes méthodologiques

Pour ces modèles trois vagues du Baromètre ont été empilées : 2015, 2016, 2017 (9070 observations).

## Bibliographie

- Duvoux Nicolas, Papuchon Adrien, « Qui se sent pauvre en France ? Pauvreté subjective et insécurité sociale », *Revue française de sociologie*, vol. 59, 2018, p. 607-647.