

# Variables Qualitatives 2 :

## Etude sur les déterminants du respect des gestes barrières et du maintien des habitudes sociales

MOUNIER TEBAS Teodoro

*Université de Nantes*

Encadré par Mme TRAVERS Muriel

---

### Abstract

Le développement suivant résume un étude réalisée au sein de l'IAE de Nantes et qui a consisté à déterminer les facteurs qui incitent les étudiants à respecter les gestes barrières depuis la fin du confinement tout en s'intéressant au phénomène de modification des habitudes sociales des étudiants durant cette même période. L'étude est construite à partir d'un sondage réalisé auprès des étudiants afin de récupérer l'information nécessaire à la construction des variables du modèle. Le langage R à été utilisé pour la manipulation des données et Stata, pour la modélisation économétrique. Le modèle retenu est le modèle 'partiel'. Par contre, le modèle ne permettra pas de prédire la modalité (1,0) correspondant aux étudiants qui respectent les gestes barrières mais n'ont pas vu leurs habitudes sociales changer.

*Keywords:* Biprobit, Biprobit récursif, Stepwise, Covid-19.

---



### 1. Introduction

Depuis mars 2020 une crise sanitaire sans précédent a touché le monde entier provoquant la mort de milliers de personnes principalement âgées mais pas seulement. Pour faire face à cette pandémie, des mesures jamais vues auparavant ont dû être mises en place telles que le confinement de toute une population, le port du masque obligatoire dans les lieux publics ou encore la distanciation sociale.

Depuis bientôt 9 mois notre société a vu ses habitudes sociales radicalement changer, et a pu mettre en lumière toutes les facettes de l'humanité fasse à un problème d'ordre général.

Si nous procédons à une caricature, il y a ceux qui prennent le virus très au sérieux en respectant les gestes barrières afin de participer à cet effort collectif car ils croient à un réel risque de santé pour eux ou pour un proche (gparents à risque). Il y a aussi ceux qui respectent les gestes barrières seulement par peur d'une répression

sans être convaincus de leur efficacité. D'autres encore sont persuadés qu'il y a une sur-réaction face à ce virus de la part de la société à cause de sa surmédiation. Ce qui est certain c'est que le virus a affecté toute la population mondiale, toutes les générations, toutes les catégories sociales que ce soit dans notre quotidien personnel et/ou professionnel.

Étant étudiants en économétrie et statistique à l'Université de Nantes, nous avons pu constater la divergence des avis à propos du covid 19 et les conséquences sur le respect des gestes barrières et le changement des habitudes chez les étudiants. Il nous a semblé donc essentiel d'étudier quels étaient les facteurs qui faisaient que certains étudiants respectaient consciencieusement les règles et d'autres moins voire pas du tout.

Nous avons donc réalisé une enquête auprès des étudiants de l'IAE de Nantes afin de récolter les informations nécessaires pour étudier **Quels sont les déterminants du respect des gestes barrières ainsi que la modification des habitudes sociales chez les étudiants ? Y a-t-il un lien entre les deux ?**

### 2. Questionnaire

#### 2.1. Comment a-t-il été réalisé ?

Le questionnaire comportait 23 questions réparties en quatre sections qui nous ont permis de construire deux variables à expliquer : le respect des gestes barrières qui

prenait la valeur 1 si l'individu respectait les gestes barrières et si non 0. De la même manière a été construite la variable qui prenait en compte le phénomène de changement des habitudes sociales.

Le reste des questions ont eu pour but de collecter des variables qui pouvaient expliquer ces deux phénomènes. Des questions à propos du covid 19 pour récolter des informations sur le suivi de l'actualité par les étudiants ainsi que leur avis à propos des mesures prises par le gouvernement. Des questions à propos de la famille, afin de savoir si l'étudiant continuait à fréquenter ses grands-parents, avait un membre de sa famille à risque ou appartenant au corps médical. Pour terminer, des questions à propos de l'étudiant pour savoir où il en était dans les études, quel âge il avait, son sexe, s'il était boursier, en emploi ou encore s'il était en bonne santé et enfin, s'il manifestait du stress vis-à-vis du covid.

### 2.2. Résultats obtenus

Nous avons récolté 555 réponses au total. Il a fallu donc traiter ces réponses parfois incomplètes et parfois absurdes, ce qui a réduit bien entendu le nombre d'observations réellement utilisées pour notre analyse. De ces 555 réponses, nous en avons retenu 431 après nettoyage. Notre échantillon, même si grandement réduit, correspond tout de même à environ 20 % de la population totale étudiée.

Les graphiques suivants représentent la répartition des individus pour chaque variable construite à partir de ce questionnaire.

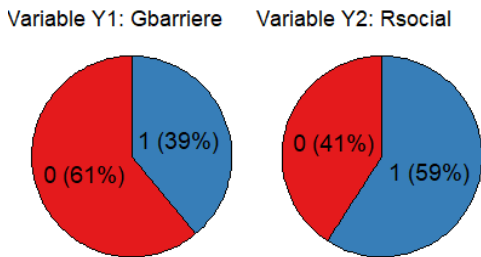


FIGURE 1: pie chart des variables à expliquer

À partir du graphique ci-dessus on peut constater que seulement 39% des étudiants respectent les gestes barrières en dehors des lieux publics où il y a obligation. Par contre 59% des étudiants ont vu leurs habitudes sociales se modifier avec le Covid. Les étudiants considérés comme respectant les gestes barrières sont ceux qui à la fois : portent leur masque en dehors des lieux publics obligatoires, respectent la distanciation sociale dans la mesure du possible et enfin se nettoient les mains au gel hydroalcoolique plus de 5 fois par jour. Seulement si ces 3 critères étaient observés alors nous les avons classés en tant que "respectant les gestes barrières".

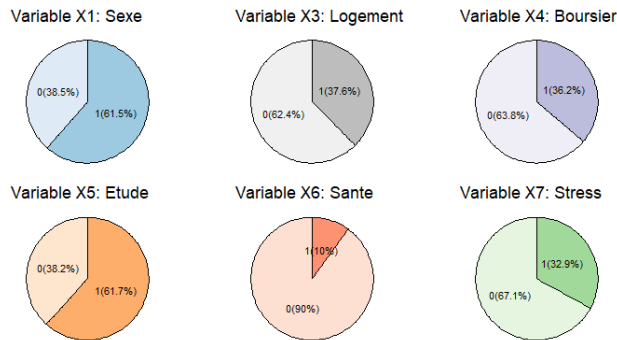


FIGURE 2: pie chart des variables à expliquer

En règle générale, nous pouvons constater que : ce sont les femmes qui ont répondu au questionnaire, les étudiants qui ne vivent pas chez leurs parents durant la semaine, qui ne sont pas boursier, qui sont en licence, en bonne santé et enfin qui ne sont pas particulièrement stressés par la situation actuelle.

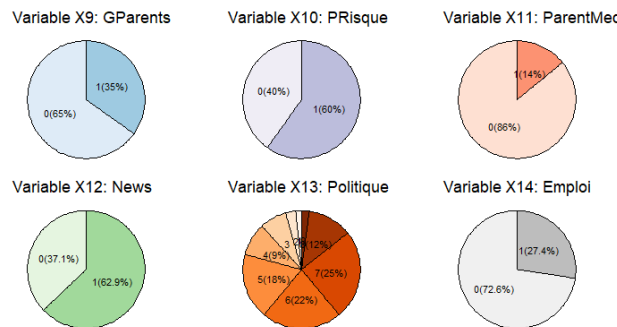


FIGURE 3: pie chart des variables à expliquer

De plus, la majorité des étudiants ayant répondu au questionnaire ne vont pas voir régulièrement leurs grands-parents, n'ont pas de parents appartenant au corps médical et ne travaillent pas en dehors de leurs études. Par contre, ils ont un parent à risque, ils suivent les informations et ils sont, de manière générale, en accord avec les mesures prises par le gouvernement actuel pour stopper cette pandémie.

Nous pouvons rajouter à ces 12 variables, la variable **Age** qui, bien entendu, ne peut être représentée en terme de répartition car nous avons décidé de la conserver telle quelle. Nous avons fait ce choix car il nous a semblé peu intéressant de construire des catégories d'âges pour des âges variants de 17 à 28 ans.

Pour terminer sur ce questionnaire, nous souhaitions savoir si le fait d'avoir été testé positif au Covid-19, pouvait avoir un impact dans le respect des gestes barrières ainsi que dans le changement des habitudes sociales. Bien entendu, la variable s'est avérée inexploitable car parmi tous les étudiants répondants, seul 1.4% d'entre eux ont eu le covid-19.

### 3. Méthodologie économétrique

#### 3.1. Un modèle Bi-probit pour estimer ces phénomènes

La méthodologie économétrique va nous permettre de répondre à notre sujet en transformant notre problématique en quelque chose qui peut se résoudre statistiquement.

Dans notre problématique nous sommes face à deux phénomènes que nous souhaitons comprendre : celui du respect des gestes barrières et celui du changement des relations sociales dû au covid. Ces deux phénomènes n'étant pas indépendants l'un de l'autre nous avons été obligés de créer un modèle qui permette de comprendre ces phénomènes ensemble. C'est ce que les modèles de type biprobit sont en mesure de réaliser. Bi, car 2 phénomènes et probit, car ce sont des phénomènes de type binaire c'est-à-dire : "peut-on y répondre par oui ou non ?"

La méthodologie bi probit nous permettra de savoir si les deux phénomènes sont vraiment liés mais aussi de comprendre quels sont les facteurs qui expliquent les 4 modalités suivantes :

- L'étudiant respecte les gestes barrières et change ses habitudes sociales. ( $Y_1 = 1; Y_2 = 1$ )
- L'étudiant ne respecte pas les gestes barrières et ne change pas ses habitudes sociales. ( $Y_1 = 0; Y_2 = 0$ )
- L'étudiant respecte les gestes barrières mais ne change pas ses habitudes sociales. ( $Y_1 = 1; Y_2 = 0$ )
- L'étudiant ne respecte pas les gestes barrières mais change ses habitudes sociales. ( $Y_1 = 0; Y_2 = 1$ )

#### 3.2. Les différentes spécifications utilisées pour l'estimation

Afin de comprendre au mieux ces deux phénomènes nous avons fait le choix de créer plusieurs modèles avec des spécifications différentes c'est-à-dire en n'incluant pas les mêmes variables explicatives pour toutes les estimations. Chaque modèle créé aura donc des hypothèses particulières en ce qui concerne les variables sélectionnées pour expliquer les variables dépendantes (gestes barrières et relations sociales).

- Modélisation avec toutes les variables explicatives que l'on appellera modèle complet.
- Modélisation avec seulement quelques variables présélectionnées (toutes les variables n'apportent pas de réponse au phénomène) que l'on appellera modèle partiel.
- Modélisation après une présélection des variables explicatives les plus pertinentes avec la variable sexe malgré sa dépendance avec Prisque. Sans la variable sexe.

- Modélisation d'un bi probit récursif c'est-à-dire que nous faisons l'hypothèse que le changement dans les relations sociales pouvait être à l'origine du respect des gestes barrières.

| Modèles<br>Variables | Complet   |         | Partiel   |         |
|----------------------|-----------|---------|-----------|---------|
|                      | GBarrière | RSocial | GBarrière | RSocial |
| Sexe                 | ✓         | ✓       | ✓         | ✓       |
| Age                  | ✓         | ✓       | ✓         | ✓       |
| Logement             |           |         |           |         |
| Boursier             | ✓         | ✓       | ✓         |         |
| Etude                | ✓         | ✓       |           | ✓       |
| Santé                | ✓         | ✓       |           | ✓       |
| Stress               |           |         |           |         |
| GParents             |           |         |           |         |
| PRisque              | ✓         | ✓       |           | ✓       |
| ParentMed            | ✓         | ✓       |           |         |
| News                 | ✓         | ✓       | ✓         | ✓       |
| Politique            | ✓         | ✓       | ✓         | ✓       |
| Emploi               |           |         |           |         |
| <b>AIC</b>           | 873.46    |         | 866.69    |         |
| <b>BIC</b>           | 954.17    |         | 924.34    |         |

TABLE 1: Tableau résumant les estimations des modèles Complet et Partiel

Le tableau précédent ainsi que celui qui suit, récapitulent l'ensemble des estimations réalisées au cours de nos recherches. Comme vous pouvez le constater, certaines variables n'ont été incluses dans aucun des modèles car elles apportaient la même information que d'autres incluses. C'est le cas de la variable *Logement*, *Stress*, *Grands Parents* et enfin *Emploi*. De plus, le modèle récursif n'a donné aucune solution car il n'a pas réussi à converger.

| Modèles<br>Variables | S.Avec<br>GB | Sexe<br>RS | Sans Sexe |    | B.Récursif |    |
|----------------------|--------------|------------|-----------|----|------------|----|
|                      | GB           | RS         | GB        | RS | GB         | RS |
| Sexe                 | ✓            | ✓          |           |    | ✓          | ✓  |
| Age                  | ✓            | ✓          | ✓         | ✓  | ✓          | ✓  |
| Logement             |              |            |           |    |            |    |
| Boursier             |              |            |           |    | ✓          |    |
| Etude                |              | ✓          |           | ✓  |            | ✓  |
| Santé                |              | ✓          |           | ✓  |            | ✓  |
| Stress               |              |            |           |    |            |    |
| GParents             |              |            |           |    |            |    |
| PRisque              |              | ✓          |           | ✓  |            | ✓  |
| ParentMed            |              |            |           |    |            |    |
| News                 | ✓            | ✓          | ✓         | ✓  | ✓          | ✓  |
| Politique            | ✓            | ✓          | ✓         | ✓  | ✓          | ✓  |
| Emploi               |              |            |           |    |            |    |
| <b>AIC</b>           | 868.50       |            | 875.39    |    | -          |    |
| <b>BIC</b>           | 922.31       |            | 921.52    |    | -          |    |

TABLE 2: Tableau résumant les estimations des modèles Stepwise et Récursif

Parmi tous les modèles estimés nous avons procédé à une sélection afin de déterminer quel modèle était le meilleur en terme de performance. Pour cela nous avons utilisé des critères d'informations sur la performance AIC<sup>1</sup> et BIC<sup>2</sup> présents dans les tableaux précédents. Sur la base du critère de l'AIC, le modèle Biprobit Partiel est celui qui le minimise et donc le modèle qu'il faudrait conserver. Si nous prenions comme critère le BIC alors, dans ce cas, ce serait le modèle "Sans Sexe" à sélectionner cette fois-ci.

Pour sélectionner un modèle, il est aussi important de comparer les qualités de prédiction des modèles. La table suivante nous présente les résultats obtenus en terme de prédictions de chaque modèles et pour chaque modalité avec les vraies valeurs observées.

| Modèle    | Modalités          | Observé | Prédit      |
|-----------|--------------------|---------|-------------|
| Complet   | $Y_1 = 1; Y_2 = 1$ | 110     | 67 (60.90%) |
|           | $Y_1 = 0; Y_2 = 0$ | 103     | 65 (63.11%) |
|           | $Y_1 = 0; Y_2 = 1$ | 101     | 28 (27.72%) |
|           | $Y_1 = 1; Y_2 = 0$ | 31      | 0 (0%)      |
| Partiel   | $Y_1 = 1; Y_2 = 1$ | 110     | 62 (56.36%) |
|           | $Y_1 = 0; Y_2 = 0$ | 103     | 61 (59.22%) |
|           | $Y_1 = 0; Y_2 = 1$ | 101     | 32 (31.68%) |
|           | $Y_1 = 1; Y_2 = 0$ | 31      | 0 (0%)      |
| Avec Sexe | $Y_1 = 1; Y_2 = 1$ | 110     | 65 (59.09%) |
|           | $Y_1 = 0; Y_2 = 0$ | 103     | 62 (60.19%) |
|           | $Y_1 = 0; Y_2 = 1$ | 101     | 28 (27.72%) |
|           | $Y_1 = 1; Y_2 = 0$ | 31      | 0 (0%)      |
| Sans Sexe | $Y_1 = 1; Y_2 = 1$ | 110     | 61 (55.45%) |
|           | $Y_1 = 0; Y_2 = 0$ | 103     | 57 (55.34%) |
|           | $Y_1 = 0; Y_2 = 1$ | 101     | 28 (27.72%) |
|           | $Y_1 = 1; Y_2 = 0$ | 31      | 0 (0%)      |

Premièrement, nous pouvons observer que la modalité ( $Y_1 = 1; Y_2 = 1$ ) c'est à dire la probabilité que l'étudiant respecte les gestes barrières et a connu un changement dans ses habitudes sociales est la modalité la mieux expliquée tout nos modèles confondus, avec un taux de bien prédit aux alentours des 2/3. De la même façon, la modalité ( $Y_1 = 0; Y_2 = 0$ ) à un taux de bien prédit avoisinant les 2/3 peu importe le modèle. Un taux de prédiction élevé pour ces deux modalités est peu surprenant car il confirme bien notre intuition du lien qu'il peut y avoir entre les habitudes sociales et le respect des gestes barrières.

En ce qui concerne les deux modalités restantes, les taux de prédiction sont nettement moins bons pour la modalité ( $Y_1 = 0; Y_2 = 1$ ) et nul pour la modalité ( $Y_1 = 1; Y_2 = 0$ ). Cela peut notamment s'expliquer par le manque

d'observations dans notre échantillon qui correspondait aux étudiants respectants les gestes barrières sans pour autant changer leurs habitudes sociales, ce qui est surprenant de manière générale.

Sur la base des critères d'informations AIC et BIC ainsi que les prédictions, le modèle Partiel est le meilleur modèle pour expliquer ces deux phénomènes conjointement.

#### 4. Conclusion

Nous pouvons conclure par l'interprétation des résultats du modèle Partiel en examinant quelles ont été les variables qui ont eu un impact significatif pour répondre à notre problématique et comment les variables explicatives ont affecté les différentes modalités.

Les principaux résultats obtenus sont les suivants : avoir un parent à risque augmente la probabilité de respecter les gestes barrières et de modifier ses fréquentations ainsi que la probabilité que l'étudiant ne respecte pas les gestes barrières mais voit ses habitudes sociales changer. Cette même variable provoque l'effet inverse lorsqu'il s'agit de la probabilité de ne pas respecter les gestes barrières et de ne pas modifier ses habitudes sociales ainsi que la probabilité de respecter les gestes barrières sans pour autant modifier ses fréquentations. Quand l'âge augmente d'un an, la probabilité de respecter les gestes barrières et de modifier ses habitudes sociales augmente d'une manière significative. A l'inverse, cela diminue de manière significative la probabilité de ne pas respecter les mesures sanitaires et de maintenir ses relations sociales.

Nous avons pu constater que ce sont les femmes, d'une manière générale, qui respectent le plus les gestes barrières et modifient leurs habitudes sociales conjointement. Tout comme les femmes, les étudiants en licence ont étonnamment plus de chances de respecter les gestes barrières tout en modifiant leurs habitudes sociales, si on compare avec les étudiants en master. Nous pourrions l'expliquer par le fait que les étudiants en licence ont souffert de changements plus importants en ce qui concerne l'enseignement notamment, ce qui est moins le cas des étudiants en master. De plus, il est intéressant d'observer qu'un étudiant boursier à une probabilité positive et significative de respecter les gestes barrières quelque soit son changement dans ses habitudes sociales, par rapport à un étudiant non boursier.

Pour terminer, un individu qui s'intéresse activement aux informations qui touchent au Covid-19 augmente sa probabilité de respecter les gestes barrières et de modifier ses habitudes par rapport à un individu qui ne le fait pas. De même, la satisfaction du répondant quant à la gestion de la crise par le gouvernement actuel est positivement corrélée avec la probabilité de respecter les gestes et de changer ses habitudes sociales.

1.  $AIC = 2\log(\mathcal{L}) - 2p$  avec  $\mathcal{L}$  le Log de Vraisemblance du modèle et  $p$  le nombre de paramètres.

2. Le BIC se calcule de manière analogue mais appliquera une pénalité plus forte au nombre de paramètres.  $AIC > BIC$  dès lors que  $\log(N) > 2$  avec  $N$  la taille de l'échantillon