

# Les déterminants du salaire au Pays-Bas

Projet d'économétrie — Département de Sciences Humaines et Sociales

École normale supérieure Paris-Saclay

Louis Bourges, Jean-Baptiste Lagrange-Dupuis et Luc Letonturier

15 mai 2023

## Introduction

Depuis Becker et sa théorie du capital humain en 1964, les travaux économiques visant à expliquer les différences de revenu entre les individus se sont multipliées. Becker a théorisé l'existence d'un calcul coût-avantage microéconomique, qui conduit les individus à arbitrer entre le coût d'une année supplémentaire d'études et le gain espéré à long terme [BECKER 1964]. Mincer, une décennie plus tard, a enrichi cette approche en incluant l'expérience accumulée au cours des années de travail dans le capital humain [MINCER 1974].

Dans notre étude, nous tenterons de mesurer les effets de ces variables mais aussi d'autres paramètres, à l'instar du genre, de la présence d'enfants, mais aussi des heures travaillées ou de l'âge. Nous nous baserons sur deux enquêtes du LISS<sup>1</sup> menées aux Pays-Bas respectivement en mai 2022 et en septembre 2022. Il s'agira, après une régression classique permettant de comprendre l'influence des différentes variables, de tester la présence d'hétéroscédasticité dans le modèle et, le cas échéant, de la corriger ; de mener un test de Chow pour tenter d'identifier d'éventuels effets de “paliers” quant au lien entre salaire et éducation ainsi que de discuter de la présence d'endogénéité dans le modèle et des moyens à notre disposition pour la corriger. Nous replacerons notre travail dans le contexte de la littérature existante et discuterons aussi de ses limites.

## 1 Présentation du modèle et de ses limites

### 1.1 Présentation des variables utilisées

Nous avons sélectionné plusieurs variables au sein de l'enquête *Work and Schooling* et de la base *Background variables*. La variable `education`, issue d'un recoupement de plusieurs variables, correspond au nombre d'années de scolarité et d'études achevée (c'est-à-dire ayant conduit à l'obtention d'un diplôme), ses valeurs sont comprises entre 0 (personne n'étant jamais allée à l'école) à 22.5 (personne titulaire d'un doctorat, sachant que la scolarité débute à l'âge de 4 ans aux Pays-Bas). La variable `genre` sépare la population en deux groupes : hommes et femmes, les autres identités de genre ayant été écartées et `age` indique l'âge des enquêtés. La variable `revenu` prend en compte le revenu brut mensuel autodéclaré, que nous avons préféré au revenu net, plus dépendant des politiques fiscales et de redistribution. La variable `heures` correspond au nombre d'heures de travail effectuées en moyenne chaque semaine tandis qu'`experience` mesure l'ancienneté des salariés dans leur entreprise (en années). Enfin, `nbenfants` indique le nombre d'enfants présents dans le foyer.

---

1. *Longitudinal Internet studies for the Social Sciences*, les questionnaires sont administrées par Centerdata

## 1.2 Détection et correction de l'hétéroscédasticité

```
## Loading required package: zoo
##
## Attaching package: 'zoo'
## The following objects are masked from 'package:base':
##
##      as.Date, as.Date.numeric
```

Afin de vérifier la présence ou non d'hétéroscédasticité au sein de notre modèle, nous avons réalisé les deux variantes du test de Breusch-Pagan (avec un test de Fisher et avec un test du rapport de vraisemblance) ainsi qu'un test de White. Tous concordent et corroborent la présence d'hétéroscédasticité, qui est par ailleurs observable graphiquement : la répartition des résidus en fonction des données prédites n'est pas homogène et l'on observe une forte variabilité de ces résidus en fonction de certaines variables du modèle, notamment la variable heures (FIGURE 1). Avec le test de Breusch-Pagan avec le rapport de vraisemblance, dont nous avons vérifié les résultats via la commande `bptest()`, nous obtenons une statistique de 75.6 et une p.value de  $2.9 \times 10^{-14}$ , qui nous permet de rejeter l'hypothèse nulle au seuil significatif de 0,01.

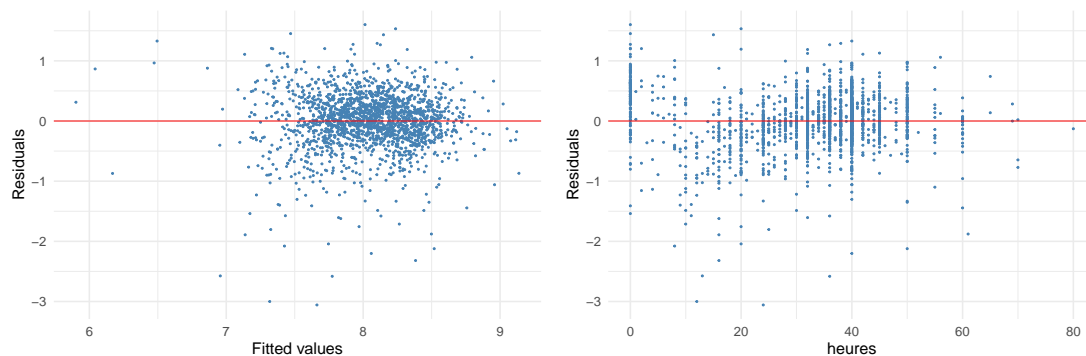


FIGURE 1 – Répartition des résidus en fonction des valeurs prédites et de la variable heures

## 1.3 Détection de l'endogénéité et pistes de correction

## 2 Résultats principaux

## 3 Analyse et mise en perspective des résultats

Et ici on peut écrire ... et insérer des blocs de code qui s'exécutent, avec le code et le résultat qui s'affichent

```
a <- 2+2
a
## [1] 4
```

ou juste le résultat :

```
## [1] 6
```

ou totalement invisibles :

Et ensuite on peut citer les résultats : à première vue  $4 < 6$  mais je crois que c'est 8 qui est le plus grand.

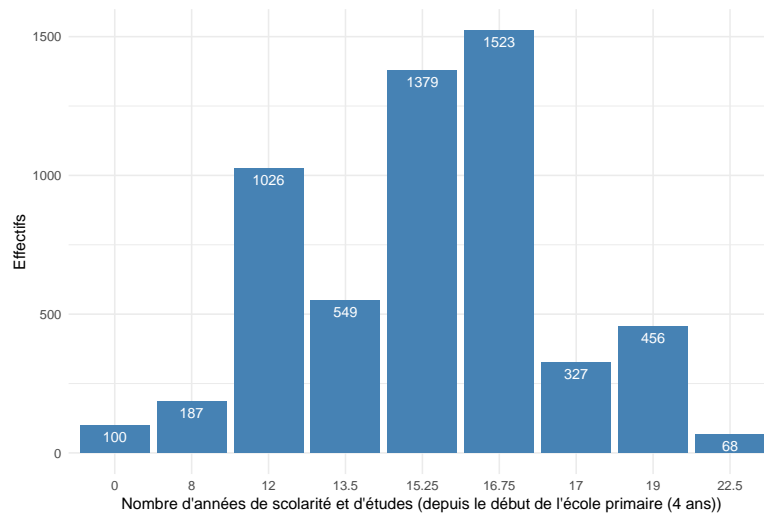


FIGURE 2 – Niveau d'éducation (avec diplôme) des individus de l'échantillon

```
## [1] 5615
## [1] 0
```

TABLE 1 – Tableau des résidus

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	5.8617	0.0959	61.13	0.0000
data\$age	0.0078	0.0010	7.77	0.0000
data\$genre	-0.3036	0.0222	-13.70	0.0000
data\$heures	0.0140	0.0008	17.01	0.0000
data\$experience	0.0030	0.0011	2.65	0.0081
data\$nbenfants	-0.0072	0.0095	-0.76	0.4483
data\$education	0.0922	0.0046	19.93	0.0000

## Références

- [1] Gary S. BECKER. « Human capital : a theoretical and empirical analysis, with special reference to education ». In : *Chicago : University of Press* 19.6 (1964), p. 4.
- [2] Jacob MINCER. « Schooling, Experience, and Earnings. » In : *Human Behavior & Social Institutions* 2 (1974).