

UNIVERSITÉ DE GRENOBLE ALPES  
(L3 MIASH, S2)

## **ÉCONOMÉTRIE**

### **TRAVAUX**

### **TRAVAIL 3**

(CETTE VERSION : 3 FÉVRIER 2025)

MICHAL W. URDANIVIA <sup>1</sup>

---

1. Contact : [michal.wong-urdanivia@univ-grenoble-alpes.fr](mailto:michal.wong-urdanivia@univ-grenoble-alpes.fr), Université de Grenoble Alpes, Faculté d'Économie, GAEL.

## TABLE DES MATIÈRES

1. Une 1ère illustration de la méthode des VIs	2
2. Une application des VIs(partie 1) :Card (1993)	2
Références	3

## 1. UNE 1ÈRE ILLUSTRATION DE LA MÉTHODE DES VIs

Soit  $Y$  une mesure de l'état de santé des personnes (cette variable est croissante avec la qualité de la santé), et  $X$  une variable indicatrice du fait de fumer (i.e.,  $X = 1$  si la personne fume et 0 sinon). Nous voulons estimer le modèle,

$$Y = \beta_1 + \beta_2 X + U$$

où  $U$  est le terme d'erreur du modèle. Pour cela, on dispose de données sur  $Y$  et  $X$ , à savoir un échantillon i.i.d  $\{(Y_i, X_i)\}_{i=1}^n$  relatives à une certaine ville. A notre grande surprise on trouve que dans nos données  $Y$  et  $X$  sont positivement corrélées. Quelques années auparavant une grande compagnie de tabac a retenu de manière aléatoire des personnes de la ville pour participer à une campagne publicitaire. La compagnie leur a offert 100 paquets de cigarettes gratuitement. Nous observons une indicatrice  $Z$  du fait d'avoir été retenu dans la campagne.

La taille totale de l'échantillon est  $n = 70$ . Nous observons  $n_{00} = 30$  individus avec  $X = Z = 0$ , et la moyenne de  $Y$  dans ce groupe est observée égale à  $\bar{Y}_{00} = 1$ . Nous observons  $n_{01} = 10$  individus avec  $X = 0$  et  $Z = 1$ , et la moyenne de  $Y$  dans ce groupe est observée égale à  $\bar{Y}_{01} = 0.8$ . Nous observons  $n_{10} = 20$  individus avec  $X = 1$  et  $Z = 0$ , et la moyenne de  $Y_i$  dans ce groupe est observée égale à  $\bar{Y}_{10} = 1.5$ . Enfin, nous observons  $n_{11}$  individus avec  $X = Z = 1$ , et la moyenne de  $Y$  dans ce groupe est observée égale à  $\bar{Y}_{11} = 1.2$ .

- (1) Pensez vous que l'hypothèse  $\mathbb{E}(U|X) = 0$  soit vraisemblable ? (Justifiez votre réponse)
- (2) Calculez l'estimateur de MCO de  $\beta_1$  et de  $\beta_2$ .
- (3) Pourquoi  $X$  peut être endogène au modèle ?
- (4) Pensez vous que  $Z$  puisse être utilisée comme variable instrumentale de  $X$  ?
- (5) Calculez l'estimateur de VI des paramètres.

## 2. UNE APPLICATION DES VIs (PARTIE 1) : CARD (1993)

Les questions suivantes ont pour but que vous précisiez plusieurs points de l'article ci-dessus en vous appuyant notamment sur votre connaissance sur le modèle de régression linéaire vu en cours.

- (1) Résumez la problématique de l'article.
- (2) Dans la spécification la plus complète (correspondant à la colonne 5 du tableau 2) que contient le vecteur  $X_i$  dans le modèle,

$$Y_i = X_i^\top \alpha + S_i \beta + U_i, \quad \mathbb{E}[X_i U_i] = 0$$

où  $S_i$  est la mesure des études,  $Y_i$  celle du salaire,  $U_i$  un terme d'erreur.

- (3) Quelle hypothèse doit-on imposer pour que l'estimateur des MCO de  $\beta$  soit sans biais.
- (4) Dans l'article une équation pour  $S_i$  est spécifiée avec,

$$S_i = X_i^\top \gamma + V_i, \quad \mathbb{E}[X_i V_i] = 0$$

où  $V_i$  est le terme d'erreur. Pourquoi il est indiqué que la corrélation entre  $U_i$  et  $V_i$  est source de biais pour l'estimateur de MCO ?

- (5) Qu'est-ce que dans l'article l'"ability biais" (voir page 9) ?
- (6) Quelle autre source de biais possible est indiqué ?
- (7) Indiquez les données employées.

## RÉFÉRENCES

Card, David. 1993. “Using geographic variation in college proximity to estimate the return to schooling.” Tech. rep., National Bureau of Economic Research. URL <http://www.nber.org/papers/w4483>.