

Econométrie des données de Panel : Exercice d'application

On se propose d'étudier l'impact de la productivité du travail, de l'usage des TIC (Technologies de l'information et de la communication) par les employés et du salaire sur le chiffre d'affaire des banques commerciales des pays de la région méditerranéenne. Pour cela, on dispose d'un panel cylindré comportant 200 banques commerciales hétérogènes observées durant la période 2006-2022.

Le modèle à estimer est le suivant :

$$(1) \quad CA_{it} = \beta_{0i} + \beta_1 PROD_{it} + \beta_2 TIC_{it} + \beta_3 SAL_{it} + \varepsilon_{it} \quad ; \quad i = 1 \dots 200 \quad ; \quad t = 2006 \dots 2022$$

Où : CA_{it} désigne le chiffre d'affaire de la banque i à la période t . $PROD_{it}$ est la productivité de travail de la banque i pour l'année t . TIC_{it} mesure le pourcentage des employés de la banque utilisant l'internet. SAL_{it} est le salaire moyen. β_{0i} désigne un effet fixe spécifique à chaque banque.

Les erreurs sont indépendantes et identiquement distribuées : $\varepsilon_{it} \rightarrow N(0, \sigma_\varepsilon^2)$.

1°a- Fournir une interprétation économique des paramètres du modèle (1).

1°b- De quel modèle s'agit-il ? Quelle est la dimension du panel utilisée dans le cadre de ce modèle ? Indiquer le nombre de degré de liberté à utiliser ?

1°c- Rappeler brièvement la procédure d'estimation du modèle (1).

1°d- Montrer analytiquement que le modèle (1) peut s'écrire sous la forme d'un modèle avec variables indicatrices individuelles.

2°- Dans ce modèle (1), on introduit une variable indicatrice relative au degré de développement du pays auquel la banque appartient. DEV_i est une variable dummy invariante dans le temps qui vaut 1 si la banque i appartient à un pays développé, 0 si la banque i appartient à un pays en développement. Le modèle à estimer devient le suivant :

$$(2) \quad CA_{it} = \beta_{0i} + \beta_1 PROD_{it} + \beta_2 TIC_{it} + \beta_3 SAL_{it} + \theta DEV_i + \varepsilon_{it}$$

- 2°a- Indiquer la banque de référence.
- 2°b- Donner une interprétation statistique du paramètre θ .
- 2°c- Interpréter économiquement le paramètre θ et préciser son signe attendu.
- 2°d- Peut-on identifier le paramètre θ ? Justifier la réponse.

3°- On considère maintenant le modèle suivant :

$$(3) \quad CA_{it} = \beta_0 + \beta_1 PROD_{it} + \beta_2 TIC_{it} + \beta_3 SAL_{it} + \varepsilon_{it} \quad ; \quad i = 1 \dots 200 \quad ; \quad t = 2006 \dots 2022$$

$$\text{Avec : } \varepsilon_{it} = u_i + v_{it} \quad ; \quad u_i \rightarrow N(0, \sigma_u^2) \quad ; \quad v_{it} \rightarrow N(0, \sigma_v^2) \quad ; \quad cov(u_i, v_{it}) = 0$$

- 3°a- Comment se fait le contrôle de l'hétérogénéité dans le modèle (3). De quel modèle s'agit-il ?
 - 3°b- Rappeler brièvement les étapes d'estimation de ce modèle.
 - 3°c- Montrer que ce modèle exploite à la fois les deux dimensions du panel.
- 4°- Après estimation des modèles (1) et (3) à l'aide d'un logiciel économétrique, on effectue le test de spécification de Hausman. Le logiciel nous fournit un p-value égal à 0,14.
- 4°a- Rappeler le principe de ce test en précisant ses hypothèses, la statistique utilisée et sa loi.
 - 4°b- Quel modèle doit-on choisir au seuil de 5% ? Justifier la réponse.