# Introduction à l'économétrie

# Tests usuels sur le modèle linéaire multiple

*Ce cours vous est proposé par Olivier Baron, Maître de conférences, Université de Bordeaux et par AUNEGe, l'Université Numérique en Économie Gestion.*

## Table des matières

# Préambule

**Objectifs :**

- Présentation de l'ensemble des tests usuels mis en œuvre sur le modèle linéaire multiple.

# Introduction

Toute estimation doit être complétée par la mise en œuvre de tests. En effet, les résultats obtenus sont des réalisations de variables aléatoires. Ils ne nous renseignent pas avec certitude sur la valeur des coefficients $\beta_j$ contribuant à générer les réalisations $y_i$ dans le cadre du modèle

$$\underset{(n,1)}{y} = \underset{(n,k)}{X} . \underset{(k,1)}{\beta} + \underset{(n,1)}{\varepsilon} .$$

Dans l'exemple présenté à la fin du premier chapitre, on a estimé une équation de salaire et trouvé une influence négative du taux de chômage sur la croissance des salaires nominaux, soit : $\hat{b}_1 = -0{,}0975$. Ici, cette valeur est une réalisation qui dépend des observations particulières des variables explicatives et expliquée figurant dans l'échantillon utilisé. On trouverait forcément un résultat différent en utilisant un autre échantillon (par exemple, des observations sur une période légèrement différente avec quelques trimestres en plus ou en moins).

On sait que les MCO sont sans biais dès que le modèle vérifie les hypothèses $H_1$, $H_2$ et $H_3$. Mais cela ne veut pas dire que $-0{,}0975 = b_1$. On sait seulement que cette valeur est la réalisation d'un estimateur $\hat{b}_1$ tel que $E(\hat{b}_1) = b_1$.

Si l'on veut définir une équation de salaire pour un modèle macroéconomique, il faut prendre des décisions, par exemple celle d'inclure ou non le taux de chômage comme variable explicative de la croissance des salaires. La question posée est alors la suivante :

le résultat $\hat{b}_1 = -0{,}0975$ correspond-il à $b_1 = 0$ ? Autrement dit, la valeur $-0{,}0975$ est-elle la réalisation d'une variable aléatoire d'espérance nulle ?

Les tests qui seront étudiés dans ce chapitre permettent de répondre à ce type de questions.

On peut distinguer deux grandes catégories de tests :

- **Les tests de restrictions sur le modèle**. Ces tests se situent à l'intérieur du cadre défini par l'écriture du modèle ( $\underset{(n,1)}{y} = \underset{(n,k)}{X} . \underset{(k,1)}{\beta} + \underset{(n,1)}{\varepsilon}$) et les hypothèses $H_1$, $H_2$, $H_3$, $H_4$ et $H_6$. Dans ce cadre, on examine dans quelle mesure il est possible de simplifier le modèle, par exemple en supprimant des variables explicatives ou en imposant des contraintes sur les coefficients.

- **Les tests de spécification**. Ces tests examinent la pertinence de la spécification définie par l'écriture du modèle et les hypothèses concernées.

En pratique une démarche rigoureuse consiste évidemment à appliquer les tests de spécification avant les tests de restrictions sur le modèle. **C'est l'ordre inverse qui sera adopté dans ce chapitre,** choix qui relève d'un objectif exclusivement pédagogique.

# Résultats préliminaires

On considère le modèle linéaire multiple :

$$\underset{(n,1)}{y} = \underset{(n,k)}{X} . \underset{(k,1)}{\beta} + \underset{(n,1)}{\varepsilon}$$

Les tests usuels reposent sur l'hypothèse de normalité des perturbations. Tout au long de ce chapitre nous adopterons les hypothèses suivantes :

. $H_1$ : $\mathrm{E}\left(\underset{(n,1)}{\varepsilon}\right) = \underset{(n,1)}{0}$

. $H_2$ : La matrice des variables explicatives $\underset{(n,k)}{X}$ est certaine.

. $H_3$ : La matrice $\underset{(n,k)}{X}$ est de rang égal à *k*.

. $H_4$ : $E\left(\underset{(n,n)}{\varepsilon . \varepsilon'}\right) = \sigma^2 . I_n$

. H6 : $\underset{(n,1)}{\varepsilon} \rightarrow N(0, \sigma^2 . I_n)$

L'hypothèse H6 implique que les perturbations $\varepsilon_i$ et $\varepsilon_j$ sont indépendantes et que les hypothèses H2 et H4 sont vérifiées. En toute rigueur, l'ensemble des hypothèses retenues est donc constitué de H2, H3 et H6.

Sous ces hypothèses nous avons établi au cours du chapitre précédent les résultats suivants :

. $\underset{(k,1)}{\hat{\beta}} \rightarrow N\left(\underset{(k,1)}{\beta}, \sigma^2 . \underset{(k,k)}{(X'X)^{-1}}\right)$ et en particulier $\hat{\beta}_j \rightarrow N(\beta_j, \sigma^2 . x^{jj})$

. $\hat{\sigma}^2 = \frac{SCR}{n-k}$ est un estimateur sans biais de $\sigma^2$ avec $(n-k) . \frac{\hat{\sigma}^2}{\sigma^2} \rightarrow \chi^2_{n-k}$

Au cours de ce chapitre, nous utiliserons la **propriété supplémentaire suivante** :

sous les hypothèses H2, H3 et H6 les variables aléatoires $\underset{(k,1)}{\hat{\beta}}$ et $\hat{\sigma}^2$ sont indépendantes.

## Démonstration en vidéo de la propriété supplémentaire

# Procédure des tests statistiques

## Les hypothèses

On distingue deux types d'hypothèses :

- Les **hypothèses simples** ($\beta_j = 2$ ; $\sigma^2 = 16$)
- Les **hypothèses multiples** ($\beta_j \neq 2$ ; $\sigma^2 > 16$)

On teste toujours une hypothèse contre une autre. L'hypothèse testée est appelée l'hypothèse nulle et on la note $H_0$. L'hypothèse contre laquelle on teste l'hypothèse nulle est appelée l'hypothèse alternative. On la note $\overline{H}$.

**À noter**

La règle suivante doit toujours être respectée : **en $H_0$, on a toujours une hypothèse simple** (on dit qu'on ne peut tester qu'une hypothèse simple).

Lors des tests présentés dans ce chapitre, **les hypothèses alternatives seront toujours des hypothèses multiples**. On pourra donc rencontrer deux cas de figures :

- Des tests bilatéraux : $$\begin{aligned} H_0 &: \beta_j = b_j \\ contre\ \overline{H} &: \beta_j \neq b_j \end{aligned}$$

- Des tests unilatéraux : $$\begin{aligned} H_0 &: \beta_j = b_j \\ contre\ \overline{H} &: \beta_j > b_j\ (ou\ \beta_j < b_j) \end{aligned}$$

## Mise en œuvre du test statistique

Une fois l'hypothèse nulle et l'hypothèse alternative correctement posées, le test statistique se conduit selon la procédure suivante :

- Définir précisément la statistique à utiliser pour effectuer le test et déterminer sa distribution de probabilité, **<u>si l'hypothèse nulle est vraie</u>** ;
- Prendre la décision en comparant la valeur critique de cette statistique lue sur la table et la valeur calculée de cette même statistique observée dans l'échantillon.

## Analyse de la décision prise

La mise en œuvre d'un test peut conduire à retenir une hypothèse conforme à la réalité. Mais on peut aussi commettre deux types d'erreurs. Les différents cas possibles peuvent être présentés à l'aide du tableau suivant :

| | Réalité | | |
|---|---|---|---|
| Décision | | $H_0$ | $\overline{H}$ |
| | $H_0$ | Pas d'erreur | Erreur de seconde espèce |
| | $\overline{H}$ | Erreur de première espèce | Pas d'erreur |

*Tableau 1 : analyse de la décision prise*

L'erreur de première espèce consiste à retenir l'hypothèse alternative alors que dans la réalité c'est l'hypothèse nulle qui prévaut. Autrement dit, on rejette à tort l'hypothèse nulle.

A l'inverse, l'erreur de seconde espèce consiste à accepter à tort l'hypothèse nulle.

Soit $\alpha$ la probabilité de commettre une erreur de première espèce, $\alpha$ est appelé risque de première espèce. Soit $\beta$ la probabilité de commettre une erreur de seconde espèce, $\beta$ est appelé risque de seconde espèce. Il est clair que le test idéal minimiserait à la fois $\alpha$ et $\beta$.

Cependant, on montre que pour un échantillon donné, minimiser $\alpha$ tend à augmenter $\beta$ et inversement. La procédure habituelle consiste alors à se fixer $\alpha$, à un niveau que l'on juge acceptable, et à minimiser pour cette valeur du risque de première espèce, le risque $\beta$. Cela revient à maximiser, pour une valeur fixée de $\alpha$, la valeur de $(1-\beta)$, appelée **puissance du test**.

Pour un risque $\alpha$, un test est dit **sans biais**, si la probabilité d'accepter l'hypothèse alternative est plus grande quand $\overline{H}$ est vraie que quand $\overline{H}$ n'est pas vraie.

Pour un risque $\alpha$, un test est dit **uniformément plus puissant sans biais**, s'il est sans biais et plus puissant que tout autre test au seuil $\alpha$.

# Tests sur la valeur d'un coefficient $\beta_j$ – Tests de significativité

On va d'abord s'intéresser aux tests portant sur la valeur d'une composante $\beta_j$ du vecteur $\underset{(k,1)}{\beta}$ avant d'examiner les tests qui portent sur la valeur du vecteur $\underset{(k,1)}{\beta}$.

Pour revenir à l'équation de salaire dont l'estimation a été présentée à la fin du chapitre précédent, on se demandait si le taux de chômage avait une réelle influence sur l'évolution des salaires nominaux. Concernant la valeur du coefficient $b_1$ du taux de chômage, cela revient à se poser la question suivante : $b_1 = 0$ ?

## Terminologie

Quand on se pose la question de la nullité d'une composante $\beta_j$ du vecteur $\underset{(k,1)}{\beta}$, on dit qu'on teste la significativité du coefficient $\beta_j$. Si, à l'issue du test, on conclut que $\beta_j = 0$, on dit que le coefficient $\beta_j$ est non significatif ou encore qu'il est non significativement différent de 0. Dans ce cas, la variable explicative $x_j$, dont $\beta_j$ est le coefficient, n'influence pas significativement la variable dépendante *y*.

De façon générale, la méthode que nous allons étudier permet de procéder à des tests sur n'importe quelle valeur de la composante $\beta_j$. On peut ainsi tester :

$$\begin{aligned} H_0 \;:&\; \beta_j = b \\ contre\ \overline{H} \;:&\; \beta_j \neq b \end{aligned}$$

où *b* est une valeur connue, fixée par celui qui met en œuvre le test.

## Statistique du test et intervalle de confiance

On sait que sous les hypothèses $H_2$, $H_3$ et $H_6$ :

$$\hat{\beta}_j \rightarrow N\left(\beta_j, \sigma^2.x^{jj}\right)$$

On en déduit en centrant et réduisant que :

$$\frac{\hat{\beta}_j - \beta_j}{\sqrt{\sigma^2.x^{jj}}} \rightarrow N(0,1)$$

À partir d'un intervalle de confiance de niveau $(1-\alpha)$ pour la loi normale centrée réduite, on peut en déduire un intervalle de confiance pour la valeur inconnue de $\beta_j$ . En notant $[-u_\alpha; u_\alpha]$ l'intervalle de confiance de niveau $(1-\alpha)$ pour la loi normale centrée réduite, on a :

$$Prob\{-u_\alpha \leq N(0,1) \leq u_\alpha\} = 1-\alpha$$

$$\Leftrightarrow \qquad Prob\left\{-u_\alpha \leq \frac{\hat{\beta}_j - \beta_j}{\sqrt{\sigma^2 . x^{jj}}} \leq u_\alpha\right\} = 1-\alpha$$

$$\Leftrightarrow \qquad Prob\{\hat{\beta}_j - u_\alpha\sqrt{\sigma^2 . x^{jj}} \leq \beta_j \ \leq \hat{\beta}_j + u_\alpha\sqrt{\sigma^2 . x^{jj}}\} = 1-\alpha$$

Dans la pratique toutefois, on ne peut pas calculer les bornes de cet intervalle de confiance puisqu'elles dépendent du paramètre inconnu $\sigma^2$.

On va donc devoir utiliser $\hat{\sigma}^2$ qui, sous les hypothèses retenues, est un estimateur sans biais de $\sigma^2$. On peut démontrer la propriété suivante :

**<u>Propriété</u>** : Sous les hypothèses $H_2$, $H_3$ et $H_6$ la statistique :

$\frac{\hat{\beta}_j - \beta_j}{\sqrt{\hat{\sigma}^2 . x^{jj}}}$ suit une loi de Student[1] à $(n-k)$ degrés de liberté.

Notons maintenant $\sigma(\hat{\beta}_j)$ l'écart-type de $\hat{\beta}_j$ $\left(\sigma(\hat{\beta}_j) = \sqrt{\sigma^2 . x^{jj}}\right)$ et $\hat{\sigma}(\hat{\beta}_j)$ l'écart-type estimé de $\hat{\beta}_j$ $\left(\hat{\sigma}(\hat{\beta}_j) = \sqrt{\hat{\sigma}^2 . x^{jj}}\right)$, la propriété précédente peut être reformulée de la façon suivante :

Sous les hypothèses $H_2$, $H_3$ et $H_6$ : $\qquad \frac{\hat{\beta}_j - \beta_j}{\hat{\sigma}(\hat{\beta}_j)} \to T_{n-k} \qquad (13)$

On constate que le fait d'estimer $\sigma^2$ nous a fait passer d'une loi normale centrée-réduite à une loi de Student. Cette dernière loi a une densité de probabilité symétrique comme la loi normale, mais plus plate et avec des queues de distribution plus épaisses. De plus, la distribution limite d'une loi de Student dont le nombre de degrés de liberté tend vers l'infini est une loi normale centrée-réduite. Dans la pratique, on identifie ces deux lois dès lors que $n-k > 120$.

En notant $[-t_\alpha; t_\alpha]$ l'intervalle de confiance de niveau $(1-\alpha)$ pour la loi de Student considérée, on en déduit pour $\beta_j$ l'intervalle de confiance suivant :

$$Prob\{\hat{\beta}_j - t_\alpha\hat{\sigma}(\hat{\beta}_j) \leq \beta_j \ \leq \hat{\beta}_j + t_\alpha\hat{\sigma}(\hat{\beta}_j)\} = 1-\alpha$$

---

[1] Par définition, une loi de Student est le rapport entre une loi normale centrée-réduite et la racine carrée d'une variable du Khi-deux divisée par son degré de liberté. Le degré de liberté de la loi de Student est égal au degré de liberté de la variable du Khi-deux qui la compose. On note cette variable $T_{n-k}$.

Les bornes de cet intervalle de confiance sont facilement calculables puisqu'elles sont fonction des estimations obtenues.

**Démonstration en vidéo de la loi de Student**


## Test sur la valeur de $\beta_j$

<u>Le test se présente de la façon suivante</u> :

$$\begin{aligned} H_0 &: \ \beta_j = b \\ contre\ \overline{H} &: \ \beta_j \neq b \end{aligned}$$

On se fixe une valeur pour le risque de première espèce $\alpha$ (appelé seuil de significativité dans le cas des tests de significativité.

La pratique la plus courante consiste à retenir $\alpha = 5\%$ ou $\alpha = 1\%$.

Sous les hypothèses retenues, on vient de voir que :

$$\frac{\hat{\beta}_j - \beta_j}{\hat{\sigma}(\hat{\beta}_j)} \to T_{n-k}$$

<u>On en déduit que</u> :

**Si $H_0$ est vraie alors** : $$\frac{\hat{\beta}_j - b}{\hat{\sigma}(\hat{\beta}_j)} \to T_{n-k} \qquad (14)$$

Le test est basé sur ce dernier résultat : on rejettera l'hypothèse nulle si la valeur prise par la statistique $\frac{\hat{\beta}_j - b}{\hat{\sigma}(\hat{\beta}_j)}$ ne correspond pas aux réalisations les plus probables d'une loi de Student à $(n-k)$ degrés de liberté.

Soit $t_\alpha^*$ la valeur, telle que la valeur absolue d'une loi de Student à $(n-k)$ degrés de liberté, n'ait qu'une probabilité a de la dépasser. $t_\alpha^*$ est appelée **valeur critique**[2]. $t_\alpha^*$ est telle que :

$$Prob\{|T_{n-k}| > t_\alpha^*\} = \alpha \text{ ou encore}$$

$$Prob\{-t_\alpha^* \leq T_{n-k} \leq t_\alpha^*\} = 1 - \alpha$$

**Décision** : Si $\left|\frac{\hat{\beta}_j - b}{\hat{\sigma}(\hat{\beta}_j)}\right| > t_\alpha^*$, on rejette l'hypothèse nulle (avec un risque de première espèce égal à a). A l'inverse, si $\left|\frac{\hat{\beta}_j - b}{\hat{\sigma}(\hat{\beta}_j)}\right| < t_\alpha^*$, on ne rejette pas l'hypothèse nulle et on considère que $\beta_j = b$.

**Explication de cette règle de décision** : supposons que l'on se fixe une valeur de $\alpha$ très faible, par exemple 0.0001. Il y a donc 9999 chances sur 10000 pour qu'une loi de Student à $(n-k)$ degrés de liberté prenne une valeur entre $-t_{0.0001}^*$ et $t_{0.0001}^*$ et donc seulement 1 chance sur 10000 qu'elle prenne une valeur en dehors de cet intervalle. Supposons que la valeur calculée de la statistique $\frac{\hat{\beta}_j - b}{\hat{\sigma}(\hat{\beta}_j)}$ soit en dehors de l'intervalle. On a dans cette situation deux façons possibles de réagir. La première consiste à dire que l'on est en train d'observer un évènement extrêmement rare. Il y avait 1 chance sur 10000 pour que cela arrive et cela vient de se produire. Nous n'avons pas eu de chance !!! Il faut bien comprendre qu'en réagissant ainsi, nous avons 1 chance sur 10000 d'avoir raison (la fois où cet évènement se produit) et donc 9999 chances sur 10000 d'avoir tort.

L'autre façon de réagir consiste à considérer qu'un tel évènement rarissime ne doit pas se produire. Une loi de Student à $(n-k)$ degrés de liberté ne prendra pas une valeur en dehors de l'intervalle, la probabilité est trop faible. Or nous venons de calculer la statistique du test et nous sommes à l'extérieur de l'intervalle. L'unique explication plausible consiste à considérer que la statistique calculée ne suit pas une loi de Student. C'est pour cette raison que nous observons cet évènement. Mais, on sait que si $H_0$ est vraie, cette statistique suit une loi Student. Il semble donc évident de conclure que l'hypothèse $H_0$ ne peut donc pas être vraie. Nous prenons donc la décision de rejeter l'hypothèse nulle. Bien sûr, en prenant cette décision on a 1 chance sur

[2] $t_\alpha^*$ est le fractile d'ordre $(1 - \frac{\alpha}{2})$ de la loi de Student à $(n-k)$ degrés de liberté et $(1-\alpha)$ est le niveau de confiance du test correspondant.

10000 de se tromper, probabilité que la loi de Student prenne une valeur à l'extérieur de l'intervalle. Cette seconde décision nous assure donc une probabilité de 0.9999 d'avoir raison. La probabilité de se tromper de 0.0001 représente la probabilité de rejeter à tort l'hypothèse nulle ce qui correspond bien à la définition du risque de première espèce.

**Propriété** : Le test sur la valeur d'une composante $\beta_j$ du vecteur $\underset{(k,1)}{\beta}$ qui vient d'être présenté est, pour le seuil $\alpha$, uniformément plus puissant sans biais.

## Test de significativité de $\boldsymbol{\beta_j}$

Ce test est une application particulière du test précédent avec $b = 0$, soit :

$$\begin{aligned} H_0 &: \beta_j = 0 \\ contre\ \overline{H} &: \beta_j \neq 0 \end{aligned}$$

Ce test est basé sur le résultat (14) qui, dans le cas où $b = 0$, s'écrit de la façon suivante :

**Si $H_0$ est vraie alors** : $$\boldsymbol{t_j = \frac{\hat{\beta}_j}{\hat{\sigma}(\hat{\beta}_j)} \rightarrow T_{n-k}} \qquad (15)$$

On remarque que dans le cas du test de significativité de $\beta_j$, la statistique utilisée est égale au rapport du coefficient estimé à son écart-type estimé. Cette statistique est couramment appelée le **t de Student** et est notée $t_j$.

La procédure à adopter pour mener ce test est la suivante :

- **Choisir une valeur** pour le risque de première espèce $\alpha$
- **Chercher** dans une table de la loi de Student **la valeur critique** correspondante $t^*$, telle que : $Prob\{|T_{n-k}| > t^*_\alpha\} = \alpha$.
- **Prendre une décision** : Si $\left|\frac{\hat{\beta}_j}{\hat{\sigma}(\hat{\beta}_j)}\right| > t^*_\alpha$, on rejette l'hypothèse nulle. On considère donc que la variable explicative correspondante $x_j$, influence de manière significative le phénomène *y*.

Généralement, les logiciels spécialisés donnent, en plus des valeurs du paramètre estimé $\hat{\beta}_j$, de son écart-type estimé $\hat{\sigma}(\hat{\beta}_j)$ et du t de Student $t_j$, la *p-value* ou la valeur estimée du risque de première espèce[3] (c'est-à-dire la probabilité de rejeter à tort l'hypothèse nulle) associée au test

[3] Cette information nous fournit la probabilité d'obtenir, pour la statistique du test, un résultat au moins aussi « extrême » que le résultat observé, si l'hypothèse nulle est vraie. Les logiciels la notent souvent Prob > $|t|$.

de significativité de chaque composante du vecteur $\underset{(k,1)}{\beta}$. Ainsi, si la *p-value* estimée est grande, la probabilité de rejeter à tort l'hypothèse nulle est grande et on ne rejette pas $H_0$. Dans le cas contraire ou la *p-value* est faible, la probabilité de rejeter à tort l'hypothèse nulle est faible et on est alors en mesure de rejeter $H_0$. On en déduit que la seule valeur estimée de la *p-value* permet de prendre une décision. Toute la question est de savoir à partir de quel niveau, on peut considérer la *p-value* comme élevée. Comme il l'a été précisé lors du point précédent, la pratique statistique consiste à se donner un niveau nominal de seuil de risque **égal à 1% ou 5%.**

## Application à l'équation de salaire

Reprenons l'estimation de l'équation de salaire présentée à la fin du premier chapitre. Pour rappel, la spécification estimée est la suivante :

$$\dot{w}_t = b_1.tc_{t-1} + b_2.\dot{pc}_t + b_3.\dot{pc}_{t-1} + b_4.\dot{pc}_{t-2} + b_5.\dot{pc}_{t-3} + b_6.(wpq)_{t-1} + b_7.D821 + b_8.D823 + b_9 + u_t$$

On s'interroge quant à l'influence du taux de chômage ($tc_{t-1}$) sur la croissance des salaires, ce qui revient à effectuer le test suivant :

$$\begin{aligned} H_0 &: \ b_1 = 0 \\ contre\ \overline{H} &: \ b_1 \neq 0 \end{aligned}$$

Les résultats de l'estimation sont redonnés ci- dessous :

Estimation d'une équation de salaire

Variable expliquée : $\dot{w}_t$

| | |
|---|---|
| $T = 159$ | $dl = 150$ |
| $R^2 = 0,913$ | $\overline{R}^2 = 0,908$ |
| $SCR = 0,0099$ | $\hat{\sigma} = 0,0081$ |
| $f = 196,767$ | $Prob > F = 0,000$ |

| Variable | Coefficient estimé | Écart-type estimé | Valeur du $t$ | Prob > $\|t\|$ |
|---|---|---|---|---|
| $tc_{t-1}$ | $-0,0975$ | $0,0201$ | $-4,8621$ | $0,000$ |
| $\dot{p}c_t$ | $0,7515$ | $0,0837$ | $8,9735$ | $0,000$ |
| $\dot{p}c_{t-1}$ | $0,2512$ | $0,0993$ | $2,5303$ | $0,012$ |
| $\dot{p}c_{t-2}$ | $0,0497$ | $0,0948$ | $0,5239$ | $0,601$ |
| $\dot{p}c_{t-3}$ | $-0,1283$ | $0,0784$ | $-1,6364$ | $0,104$ |
| $(wpq)_{t-1}$ | $-0,0494$ | $0,0169$ | $-2,9284$ | $0,004$ |
| $D821$ | $0,0007$ | $0,0082$ | $0,0882$ | $0,930$ |
| $D823$ | $-0,0049$ | $0,0082$ | $-0,5972$ | $0,551$ |
| Constante | $-0,2694$ | $0,0967$ | $-2,7866$ | $0,006$ |

Données trimestrielles pour l'économie française.
Période : 1951-T2 à 1990-T4.

$\dot{w}_t$ : taux de croissance du salaire nominal.
$tc_{t-1}$ : taux de chômage du trimestre précédent.
$\dot{p}c_t$ : taux de croissance des prix à la consommation du trimestre $t$.
$\dot{p}c_{t-1}$ : taux de croissance des prix à la consommation du trimestre $t-1$.
$\dot{p}c_{t-j}$ : taux de croissance des prix à la consommation du trimestre $t-j$.
$(wpq)_{t-1}$ : terme de rappel.
$D821$ : variable indicatrice du premier trimestre de 1982.
$D823$ : variable indicatrice du troisième trimestre de 1982.

*Figure 1 : estimation d'une équation de salaire*

La régression est telle que $T - k = 150 > 120$.

Dans cette situation, et compte tenu du point précédent, on déduit que :

Si H₀ est vraie alors : $t_1 = \frac{\hat{b}_1}{\hat{\sigma}(\hat{b}_1)} \rightarrow T_{150}$

Pour un risque de première espèce $\alpha = 5\%$, on trouve la valeur critique $t^*_{0,05} = 1{,}96$.

On a obtenu $\hat{b}_1 = -0{,}0975$, $\hat{\sigma}(\hat{b}_1) = 0{,}0201$ et le logiciel fournit pour le *t* de Student la valeur $t_1 = \frac{-0{,}0975}{0{,}0201} = -4{,}8621$. On constate immédiatement que $|-4{,}8621| > 1{,}96$.

On en tire la conclusion que $\frac{\hat{b}_1}{\hat{\sigma}(\hat{b}_1)}$ **ne suit pas** une loi de Student à 150 degrés de liberté et donc que $b_1$ est significativement différent de zéro. Cette conclusion est vraie au seuil $\alpha = 5\%$ mais aussi pour des valeurs de $\alpha$ beaucoup plus faibles (pour $\alpha = 0{,}001$, la table donne $t^*_{0,001} = 3{,}29$).

Sur le listing, la colonne intitulée $Prob > |t|$ fournit directement, pour chaque coefficient $b_j$, la p-value associée à la valeur de son *t* de Student $t_j$. Plus exactement, on trouve dans cette colonne, pour chaque coefficient $b_j$, les seuils de significativité $\alpha_j$ définis par :

$$\alpha_j = Prob\{|T_{150}| > t_j\}$$

En conclusion, on peut dire que les variables $\dot{pc}_{t-2}$, $\dot{pc}_{t-3}$, $D821$ et $D823$ sont non significatives et n'influencent donc pas la variable dépendante $\dot{w}_t$. Elles peuvent donc être éliminées de la liste des variables explicatives. Nous verrons par la suite que les conserver n'introduit pas de biais, mais nuit à la précision de l'estimation.

## Test sur la valeur du vecteur $\underset{(k,1)}{\beta}$

L'estimation du modèle linéaire multiple $\underset{(n,1)}{y} = \underset{(n,k)}{X} . \underset{(k,1)}{\beta} + \underset{(n,1)}{\varepsilon}$ conduit au vecteur estimé $\underset{(k,1)}{\hat{\beta}}$, à partir duquel on peut réaliser le test suivant :

$$\begin{aligned} H_0 : \ & \underset{(k,1)}{\beta} = \underset{(k,1)}{b} \\ contre\ \overline{H} : \ & \underset{(k,1)}{\beta} \neq \underset{(k,1)}{b} \end{aligned}$$

où $\underset{(k,1)}{b}$ est un vecteur défini par celui qui met en œuvre le test. Par exemple, on pourrait se demander si tous les coefficients sont nuls, soit : $\underset{(k,1)}{\beta} = \underset{(k,1)}{0}$ ?

### Remarque

Ce dernier test est présenté pour des raisons essentiellement pédagogiques. Dans la pratique, il n'est pas très intéressant car la constante n'est pas nulle en général, même si aucune variable explicative n'a d'influence sur le phénomène *y*. On verra plus loin un test plus utile, dans lequel l'hypothèse examinée est : « tous les coefficients sont nuls, sauf la constante ».

On sait que sous les hypothèses $H_2$, $H_3$ et $H_6$ on a :

$$\underset{(k,1)}{\hat{\beta}} \rightarrow N\left(\underset{(k,1)}{\beta}, \sigma^2 . \underset{(k,k)}{(X'X)^{-1}}\right)$$

De manière équivalente, on peut donc écrire que :

$$\underset{(k,1)}{\hat{\beta}} - \underset{(k,1)}{\beta} \to N\left(\underset{(k,1)}{0}, \sigma^2.\underset{(k,k)}{(X'X)^{-1}}\right)$$

En utilisant une propriété des vecteurs gaussiens[4], on en déduit que :

$$(\hat{\beta} - \beta)'[\sigma^2(X'X)^{-1}]^{-1}(\hat{\beta} - \beta) \to \chi_k^2$$

Ou encore :

$$(\hat{\beta} - \beta)'\frac{X'X}{\sigma^2}(\hat{\beta} - \beta) \to \chi_k^2 \qquad (16)$$

De la même façon que pour les tests de Student examinés dans la section précédente, la statistique apparaissant dans l'expression (16) ne peut pas être utilisée directement pour un test sur la valeur de $\underset{(k,1)}{\beta}$ car elle dépend de $\sigma^2$ dont la valeur est inconnue. Comme précédemment, le fait de remplacer le paramètre $\sigma^2$ par son estimation, va modifier la loi suivie par la statistique du test.

### Démonstration en vidéo de la loi Fisher


Ainsi, on peut montrer la propriété suivante :

**<u>Propriété</u>** : Sous les hypothèses $H_2$, $H_3$ et $H_6$ la statistique *f*, définie par :

---

4 Soit $\underset{(k,1)}{y}$ un vecteur colonne à *k* lignes tel que $\underset{(k,1)}{y} \to N(0\,, \Sigma)$, avec $\Sigma$ régulière, alors $y'\Sigma^{-1}y \to \chi_k^2$.

$f = (\hat{\beta} - \beta)' \frac{X'X}{k\hat{\sigma}^2}(\hat{\beta} - \beta)$ suit une loi de Fisher[5] de degrés de liberté *k* et *n-k*.

**Revenons au test de l'hypothèse** :

$$H_0 : \underset{(k,1)}{\beta} = \underset{(k,1)}{b}$$
$$contre\ \overline{H} : \underset{(k,1)}{\beta} \neq \underset{(k,1)}{b}$$

En se basant sur la propriété qui vient d'être établie, on peut effectuer le test à partir du résultat suivant :

**Si $H_0$ est vraie** alors : $f = (\hat{\beta} - b)' \frac{X'X}{k\hat{\sigma}^2}(\hat{\beta} - b) \rightarrow F(k, n-k)$ (17)

Comme nous l'avons précisé au départ, étudier le test sur la valeur du vecteur $\underset{(k,1)}{\beta}$ présente un intérêt essentiellement pédagogique. Cela permet, à partir de l'expression (17) de définir pour $\underset{(k,1)}{\beta}$ une région de confiance qui pourra être comparée avec l'ensembles des intervalles de confiance que l'on peut construire pour chacune des composantes $\beta_j$ du vecteur $\underset{(k,1)}{\beta}$.

En particulier, il est possible de conclure à partir d'un test global que $\underset{(k,1)}{\beta} \neq \underset{(k,1)}{0}$ et d'obtenir, sur le même modèle, par un test de Student effectué sur chacun des *k* coefficients, qu'aucun de ceux-ci ne sont significatifs ($\beta_j = 0 \quad \forall j = 0 \ldots (k-1)$). La différence entre les deux démarches provient du fait que l'on ne tient pas compte, dans les tests de Student, de la covariance entre les estimateurs des composantes $\beta_j$ du vecteur $\underset{(k,1)}{\beta}$.

En principe, $Cov(\hat{\beta}_i, \hat{\beta}_j) \neq 0$, puisque la matrice de variances-covariances du vecteur $\underset{(k,1)}{\hat{\beta}}$ ($\sigma^2. \underset{(k,k)}{(X'X)^{-1}}$), n'est en général pas diagonale.

Seul le test global défini par (17) tient compte des covariances entre les estimateurs des composantes de $\underset{(k,1)}{\beta}$, puisque la statistique du test est formée à partir de l'expression (16) qui intègre la matrice de variances-covariances de $\underset{(k,1)}{\hat{\beta}}$. En revanche, les statistiques utilisées dans les tests de significativité effectués coefficient par coefficient, ne tiennent compte que des variances estimées de chaque coefficient estimé.

---

[5] Par définition, une variable de Fisher correspond au rapport de deux variables de Khi-deux indépendantes, chacune divisée par son degré de liberté. Les deux degrés de liberté de la loi de Fisher sont ceux des deux Khi-deux qui la composent. Si ces degrés de liberté sont $\nu_1$ et $\nu_2$, on notera cette variable $F(\nu_1, \nu_2)$.

D'autre part, il faut aussi comprendre qu'un test global (utilisant une loi de Fisher) et un test consistant à réaliser un test de Student sur chaque coefficient, n'ont pas en général le même risque de première espèce.

Supposons que l'on retienne $\alpha = 5\%$ pour le test global et $\alpha = 5\%$ pour chacun des tests de Student. L'erreur de première espèce du test consistant à effectuer un test de Student sur chaque coefficient a une probabilité bien supérieure à 5% puisque, l'hypothèse nulle ($\beta_0 = \beta_1 = \cdots = \beta_{k-1} = 0$ sera rejetée si $\beta_0 = 0$ est rejetée par le premier test de Student (évènement de probabilité égale à 5%) OU si $\beta_1 = 0$ est rejetée par le deuxième test de Student (évènement de probabilité égale à 5%) OU ….. OU si $\beta_{k-1} = 0$ est rejetée par le $k^{\text{ième}}$ test de Student (évènement de probabilité égale à 5%).

### Remarque

L'ensemble de ces remarques montre bien qu'il est plus pertinent d'utiliser un test global (utilisant une loi de Fisher) dès que l'on veut effectuer un ensemble de tests sur la valeur des coefficients. Toutefois, tel qu'il est présenté ici, le test de Fisher ne correspond pas tout à fait aux questions que l'on se pose dans la pratique.

Considérons à nouveau le vecteur $\underset{(k,1)}{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_{k-1} \end{pmatrix}$ où $\beta_0$ est la constante.

On peut vouloir tester l'hypothèse $H_{01}$ : $[\beta_1 = -0{,}1 \text{ et } \beta_2 = 0]$ sans s'interroger sur la valeur des autres coefficients, lesquels peuvent ne pas présenter d'enjeu particulier sur le plan économique. On peut aussi vouloir tester $H_{02}$ : $[\beta_1 = \beta_2 = \cdots = \beta_{k-1} = 0]$, c'est-à-dire se demander si tous les coefficients autres que la constante sont nuls.

Le test sur la valeur de $\underset{(k,1)}{\beta}$ qui vient d'être présenté ne permet pas de tester $H_{01}$ ni $H_{02}$. En effet lorsqu'on teste $H_0$ : $\left[\underset{(k,1)}{\beta} = \underset{(k,1)}{b}\right]$, il faut donner la valeur de toutes les composantes du vecteur $\underset{(k,1)}{\beta}$. Or, dans la pratique, on n'a pas d'idée à priori sur la valeur de toutes les composantes $\beta_j$ de $\underset{(k,1)}{\beta}$. On ne s'intéresse même pas nécessairement à la valeur de chaque $\beta_j$. La constante, par exemple, n'a souvent pas de contenu théorique particulier. Si on désire tester l'hypothèse $H_{02}$ en effectuant un test sur la valeur de $\underset{(k,1)}{\beta}$ où $H_0$ est de la forme :

$$\text{H}_0 : \underset{(k,1)}{\beta} = \underset{(k,1)}{b} \text{ avec } \underset{(k,1)}{b} = \begin{pmatrix} b_0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

on prend le risque de rejeter $H_0$ uniquement parce qu'on s'est trompé sur la valeur $b_0$ de $\beta_0$, alors que $H_{02}$ est peut-être vraie.

Dans la section suivante, nous allons étudier des procédures permettant de tester des hypothèses de la forme $H_{01}$ ou $H_{02}$.

# Contraintes linéaires sur les coefficients : estimation et tests

On considère à nouveau le modèle linéaire multiple en supposant que les hypothèses $H_2$, $H_3$ et $H_6$ sont vérifiées. Jusqu'à présent, nous avons défini la procédure d'estimation en supposant que $\underset{(k,1)}{\beta}$ pouvait prendre n'importe quelle valeur dans $R^k$. Or, on peut se demander s'il n'existe pas une ou plusieurs contraintes linéaires liant les coefficients $\beta_j$ entre eux. Par exemple on pourrait se demander si : $[\beta_1 = \beta_2$ et $\beta_3 = \beta_4]$ ?

Plus généralement, on formulera l'existence de *r* contraintes linéaires indépendantes sur les *k* paramètres $\beta_j$ de la façon suivante :

$$\begin{cases} \underset{(r,k)}{C} . \underset{(k,1)}{\beta} = \underset{(r,1)}{q} \\ \text{avec } Rang \underset{(r,k)}{C} = r \text{ et } r < k \end{cases}$$

. $\underset{(r,k)}{C}$ est une matrice certaine connue de l'économètre et $\underset{(r,1)}{q}$ est un vecteur certain et connu.

. $\underset{(r,k)}{C}$ est de rang égal à *r* car on suppose que les *r* contraintes linéaires sont linéairement indépendantes.

. $r < k$ car si *r* était égal à *k*[6], la démarche consistant à effectuer une estimation serait évacuée par l'imposition des contraintes, puisqu'alors : $\underset{(k,1)}{\beta} = \underset{(k,k)}{C^{-1}} . \underset{(k,1)}{q}$

[6] Si *r* > *k* et si les *r* contraintes sont indépendantes, le système précédent n'admet aucune solution.

## Exemples

On peut se demander si :

**1.** $[\boldsymbol{\beta_1 = \beta_2}]$**, c'est-à-dire si** $[\boldsymbol{\beta_1 - \beta_2 = 0}]$**.**

L'écriture de cette contrainte sous la forme $C.\beta = q$ se présente de la façon suivante :

$$\begin{pmatrix} 0 & 1 & -1 & 0 & \cdots & 0 \end{pmatrix} . \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_{k-1} \end{pmatrix} = 0$$

avec $\underset{(1,k)}{C} = \begin{pmatrix} 0 & 1 & -1 & 0 & \cdots & 0 \end{pmatrix}$ et $\underset{(1,1)}{q} = 0$

**2.** $[\boldsymbol{\beta_1 = \beta_2} \textbf{ et } \boldsymbol{\beta_3 = 2\beta_4 - 1}]$**, c'est-à-dire si** $[\boldsymbol{\beta_1 - \beta_2 = 0} \textbf{ et } \boldsymbol{\beta_3 - 2\beta_4 = -1}]$**.**

L'écriture de ces 2 contraintes sous la forme $C.\beta = q$ se présente de la façon suivante :

$$\begin{pmatrix} 0 & 1 & -1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & -2 & \cdots & 0 \end{pmatrix} . \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \\ \vdots \\ \beta_{k-1} \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$$

avec $\underset{(2,k)}{C} = \begin{pmatrix} 0 & 1 & -1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & -2 & \cdots & 0 \end{pmatrix}$ et $\underset{(2,1)}{q} = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$

**3.** $[\boldsymbol{\beta_1 = \beta_2 = \cdots = \beta_{k-1} = 0}]$ **(tous les coefficients sont nuls, sauf la constante).**

Ces ($k$-1) contraintes s'écrivent :

$$\begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} . \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \\ \vdots \\ \vdots \\ \beta_{k-1} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

avec $\underset{(k-1,k)}{C} = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$ et $\underset{(k-1,1)}{q} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$. Ici, puisqu'aucune contrainte ne concerne la constante $\beta_0$, la première colonne de C ne contient que des 0.

On pourrait procéder à l'estimation d'un modèle en imposant au vecteur estimé de respecter les *r* contraintes $C.\beta = q$. Mais cette démarche n'est justifiée que si l'on a établi que ces contraintes correspondent à la réalité.

Dans ce qui va suivre, nous présenterons tout d'abord la méthode d'estimation sous contraintes et ses propriétés. Puis nous exposerons la procédure permettant de tester l'hypothèse $\underset{(r,k)}{C} . \underset{(k,1)}{\beta} = \underset{(r,1)}{q}$.

## Estimation d'un modèle sous contraintes

La prise en compte de contraintes dans l'estimation d'un modèle linéaire peut être effectuée de deux façons :

- en procédant à des changements de variables pour intégrer les contraintes dans la spécification du modèle.
- en appliquant directement la procédure d'estimation sous contraintes.

### Intégration des contraintes dans l'écriture du modèle

Estimer un modèle avec *k* coefficients sur lesquels on impose *r* contraintes, revient à estimer seulement $(k - r)$ coefficients. On peut donc intégrer les contraintes en procédant à des changements de variables pour réécrire le modèle en fonction de seulement $(k - r - 1)$ variables explicatives. L'application des MCO au modèle transformé correspond à l'estimation sous contraintes du modèle initial. Il faut noter que cette procédure intuitive n'est pas toujours facile à appliquer car les transformations à réaliser ne sont pas nécessairement évidentes.

**Exemple :**

Modèle sans contrainte : $y_i = \beta_0 + \beta_1.x_{1,i} + \beta_2.x_{2,i} + \beta_3.x_{3,i} + \beta_4.x_{4,i} + \beta_5.x_{5,i} + \varepsilon_i$

Nombre de variables explicatives du modèle sans contrainte : 5

Contraintes : $[\beta_1 = \beta_2 \text{ et } \beta_3 = 2\beta_4]$

Intégration des contraintes dans le modèle :

$$y_i = \beta_0 + \beta_1.x_{1,i} + \beta_1.x_{2,i} + (2\beta_4).x_{3,i} + \beta_4.x_{4,i} + \beta_5.x_{5,i} + \omega_i$$[7]

$\Leftrightarrow \quad y_i = \beta_0 + \beta_1.(x_{1,i} + x_{2,i}) + \beta_4.(2x_{3,i}+x_{4,i}) + \beta_5.x_{5,i} + \omega_i$

$\Leftrightarrow \quad y_i = \beta_0 + \beta_1.z_{1,i} + \beta_4.z_{2,i} + \beta_5.x_{5,i} + \omega_i$

avec $z_{1,i} = x_{1,i} + x_{2,i}$ et $z_{2,i} = 2x_{3,i}+x_{4,i}$

Nombre de variables explicatives du modèle contraint : 3

### Procédure directe d'estimation sous contraintes

Il est toujours possible de procéder à l'estimation d'un modèle sous contraintes. La plupart des logiciels économétriques permettent de programmer ce type d'estimation avec une ligne de commande supplémentaire. La méthode retenue est celle des MCO sous contraintes définie par :

$\hat{\beta}_c$ est solution du programme $\underset{\beta}{\mathrm{Min}} \underset{(1,n)}{\left(y - X\hat{\beta}\right)'}.\underset{(n,1)}{\left(y - X\hat{\beta}\right)}$

$$\text{sc} \quad \left| \underset{(r,k)}{C}.\underset{(k,1)}{\hat{\beta}} = \underset{(r,1)}{q} \right.$$

$\hat{\beta}_c$ est l'estimateur des MCO sous contraintes du modèle. On notera $\hat{\beta}_{nc}$ l'estimateur des MCO qui ne prend pas en compte les contraintes (auparavant noté $\hat{\beta}$ et égal à $(X'X)^{-1}X'y$).

<u>**Théorème**</u> : Soit le modèle linéaire multiple $\underset{(n,1)}{y} = \underset{(n,k)}{X}.\underset{(k,1)}{\beta} + \underset{(n,1)}{\varepsilon}$. L'estimateur du vecteur $\underset{(k,1)}{\beta}$ par les MCO sous les contraintes $\underset{(r,k)}{C}.\underset{(k,1)}{\beta} = \underset{(r,1)}{q}$ est défini par :

$$\hat{\beta}_c = \hat{\beta}_{nc} - (X'X)^{-1}C'[C(X'X)^{-1}C']^{-1}\left(C\hat{\beta}_{nc} - q\right) \qquad (18)$$

[7] On prévoit des notations différentes pour les perturbations car si les contraintes ne correspondent pas à la réalité, les erreurs n'ont pas le même contenu dans les modèles contraint et non contraint.

## Démonstration en vidéo du théorème


**Propriétés de l'estimateur contraint :**

Pour étudier les propriétés de l'estimateur défini par la relation (18) on distingue le cas où les contraintes sont pertinentes du cas où l'on commet une erreur en les imposant lors de l'estimation. Notons $H_0$ l'hypothèse que les contraintes sont vraies :

$$H_0 : \qquad \underset{(r,k)}{C}.\underset{(k,1)}{\beta} = \underset{(r,1)}{q}$$

- Premier cas : $H_0$ est vraie.

**Propriétés** : Sous les hypothèses $H_1$, $H_2$, $H_3$, $H_4$ et $H_0$, l'estimateur des MCO contraint $\hat{\beta}_c$ est sans biais et plus précis que l'estimateur non contraint $\hat{\beta}_{nc}$. Autrement dit, l'estimateur est sans biais et l'on gagne en précision lorsqu'on intègre les contraintes définies par $H_0$. Ce résultat est intuitif : les contraintes $\underset{(r,k)}{C}.\underset{(k,1)}{\beta} = \underset{(r,1)}{q}$ constituent un ensemble d'informations sur le vecteur $\underset{(k,1)}{\beta}$ dont la prise en compte réduit l'incertitude affectant l'estimation.

- Deuxième cas : $H_0$ n'est pas vérifiée $\Leftrightarrow \underset{(r,k)}{C}.\underset{(k,1)}{\beta} \neq \underset{(r,1)}{q}$

**Propriété** : Lorsque $H_0$ n'est pas vérifiée, l'estimateur contraint $\hat{\beta}_c$ est biaisé.

## Démonstration en vidéo de l'estimateur contraint sans biais ou biaisé


## Démonstration en vidéo sur la précision de l'estimateur contraint

## Test de $r$ contraintes linéaires sur les composantes du vecteur $\underset{(k,1)}{\boldsymbol{\beta}}$

**Il s'agit de tester l'hypothèse $H_0$ :** $\underset{(r,k)}{\boldsymbol{C}} . \underset{(k,1)}{\boldsymbol{\beta}} = \underset{(r,1)}{\boldsymbol{q}}$ .

Le test se présente de la façon suivante :

$$\begin{cases} H_0 : & \begin{cases} \underset{(n,1)}{y} = \underset{(n,k)}{X} . \underset{(k,1)}{\beta} + \underset{(n,1)}{\varepsilon} \\ \text{sc } \underset{(r,k)}{C} . \underset{(k,1)}{\beta} = \underset{(r,1)}{q} \end{cases} \\ contre\ \overline{H} & : \underset{(n,1)}{y} = \underset{(n,k)}{X} . \underset{(k,1)}{\beta} + \underset{(n,1)}{\varepsilon} \end{cases}$$

Autrement dit, on veut tester le modèle contraint contre le modèle non contraint.

### Notations

Comme nous l'avons déjà précisé, l'estimateur contraint sera noté $\hat{\beta}_c$ et l'estimateur non contraint $\hat{\beta}_{nc}$. Le vecteur des prédictions de la variable expliquée du modèle contraint est $\underset{(n,1)}{\hat{y}_c}$ ($\underset{(n,1)}{\hat{y}_{nc}}$ dans le cas du modèle non contraint). $\underset{(n,1)}{\hat{y}_c}$ correspond à la projection du vecteur $\underset{(n,1)}{y}$ sur un espace de dimension $(k-r)$, noté $\mathcal{L}_C(X)$. Soit $\underset{(n,1)}{\widehat{\omega}}$ le vecteur des résidus estimés de la régression contrainte. $\underset{(n,1)}{\widehat{\omega}}$ correspond à la projection du vecteur $\underset{(n,1)}{y}$ sur l'espace orthogonal à $\mathcal{L}_C(X)$, noté $\mathcal{L}_C^{\perp}(X)$ et de dimension $\big(n-(k-r)\big)$.

On appellera degrés de liberté de l'estimation contrainte, notés $dl_c$, la dimension de $\mathcal{L}_C^{\perp}(X)$ soit $dl_c = n-k+r$ (dans le cas du modèle non contraint on a $dl_{nc} = n-k$).

Soit $\underset{(1,n)(n,1)}{\widehat{\omega}' \ \widehat{\omega}}$ la somme des carrés des résidus de la régression contrainte ( $\underset{(1,n)(n,1)}{\hat{\varepsilon}' \ \hat{\varepsilon}}$ dans le cas du modèle non contraint). On notera $SCR_c = \underset{(1,n)(n,1)}{\widehat{\omega}' \ \widehat{\omega}}$ et $SCR_{nc} = \underset{(1,n)(n,1)}{\hat{\varepsilon}' \ \hat{\varepsilon}}$.

Enfin, le coefficient de détermination de la régression contrainte sera noté $R_c^2$ ($R_{nc}^2$ dans le cas du modèle non contraint).

### Définition du test

On sait que sous les hypothèses $H_2$, $H_3$ et $H_6$, $\underset{(k,1)}{\hat{\beta}_{nc}} \rightarrow N\left(\underset{(k,1)}{\beta} \ ; \ \sigma^2 . \underset{(k,k)}{(X'X)^{-1}}\right)$.

On en déduit que, puisque la matrice $\underset{(r,k)}{C}$ est une matrice certaine :

$$\underset{(r,k)}{C}.\underset{(k,1)}{\hat{\beta}_{nc}} \to N\left(\underset{(r,k)}{C}.\underset{(k,1)}{\beta}\ ;\ \sigma^2.\underset{(r,k)}{C}\underset{(k,k)}{(X'X)^{-1}}\underset{(k,r)}{C'}\right)$$ [8]

et donc que :

$$\underset{(r,k)}{C}.\underset{(k,1)}{\hat{\beta}_{nc}} - \underset{(r,k)}{C}.\underset{(k,1)}{\beta} \to N\left(\underset{(r,1)}{0}\ ;\ \sigma^2.\underset{(r,k)}{C}\underset{(k,k)}{(X'X)^{-1}}\underset{(k,r)}{C'}\right)$$

Si $H_0$ est vraie, alors $\underset{(r,k)}{C}.\underset{(k,1)}{\beta} = \underset{(r,1)}{q}$, ce qui implique que :

$$\underset{(r,k)}{C}.\underset{(k,1)}{\hat{\beta}_{nc}} - \underset{(r,1)}{q} \to N\left(\underset{(r,1)}{0}\ ;\ \sigma^2.\underset{(r,k)}{C}\underset{(k,k)}{(X'X)^{-1}}\underset{(k,r)}{C'}\right)$$

En utilisant la même propriété qui a permis d'aboutir à la relation (16), on peut en déduire un test basé sur le résultat suivant :

**Si $H_0$ est vraie** alors : $\left(C\hat{\beta}_{nc} - q\right)'[\sigma^2 C(X'X)^{-1}C']^{-1}\left(C\hat{\beta}_{nc} - q\right) \to \chi^2_r$ (19)

Dans la pratique, toutefois, on ne peut pas utiliser la statistique figurant dans l'expression précédente car $\sigma^2$ est un paramètre inconnu qui doit être estimé. On utilisera, comme nous l'avons déjà fait, l'estimateur sans biais $\hat{\sigma}^2$ défini par :

$$\hat{\sigma}^2 = \frac{SCR_{nc}}{n-k}$$

On sait que sous les hypothèses $H_2$, $H_3$ et $H_6$, $\hat{\sigma}^2$ et $\hat{\beta}_{nc}$ sont indépendants et que $(n-k).\frac{\hat{\sigma}^2}{\sigma^2} \to \chi^2_{n-k}$.

Formons le rapport :

$$\frac{\dfrac{\left(C\hat{\beta}_{nc} - q\right)'[\sigma^2 C(X'X)^{-1}C']^{-1}\left(C\hat{\beta}_{nc} - q\right)}{r}}{\dfrac{(n-k).\dfrac{\hat{\sigma}^2}{\sigma^2}}{n-k}}$$

Si $H_0$ est vraie, il s'agit du rapport de deux variables de khi-deux indépendantes, chacune divisée par son degré de liberté. Ce rapport suit donc, par définition, une loi de Fisher.

Après simplifications, on aboutit à la propriété suivante :

**Si $H_0$ est vraie** alors : $f = \frac{\left(C\hat{\beta}_{nc} - q\right)'[C(X'X)^{-1}C']^{-1}\left(C\hat{\beta}_{nc} - q\right)}{r\hat{\sigma}^2} \to F_{(r,n-k)}$ (20)

[8] $\underset{(r,r)}{\mathrm{Var}}\, C\hat{\beta}_{nc} = \mathrm{E}\left[\left(C\hat{\beta}_{nc} - C\beta\right)\left(C\hat{\beta}_{nc} - C\beta\right)'\right] = \mathrm{C.E}\left[\left(\hat{\beta}_{nc} - \beta\right)\left(\hat{\beta}_{nc} - \beta\right)'\right].C' = \sigma^2.C(X'X)^{-1}C'$

Le test de r contraintes linéaires sur les composantes du vecteur $\underset{(k,1)}{\beta}$ est basé sur le résultat (20).

On examine si la réalisation de la statistique f correspond à une valeur qu'une loi de Fisher $F_{(r,n-k)}$, n'a pas une probabilité trop faible d'atteindre.

## Écriture de la statistique du test en fonction des sommes des carrés des résidus

Au numérateur de la statistique *f* apparaissant dans la relation (20), figure une expression dont on peut montrer qu'elle est égale à la différence entre les sommes des carrés des résidus des régressions contrainte et non contrainte.

### Écriture de la statistique *f* en fonction des sommes des carrés des résidus


Compte tenu du fait que $\hat{\sigma}^2 = \frac{SCR_{nc}}{n-k}$ où $n-k = dl_{nc}$ et que $r = dl_c - dl_{nc}$[9], on en déduit la propriété suivante :

**<u>Propriété</u>** : Soit un modèle linéaire multiple de la forme $\underset{(n,1)}{y} = \underset{(n,k)}{X} . \underset{(k,1)}{\beta} + \underset{(n,1)}{\varepsilon}$, muni des hypothèses $H_2$, $H_3$ et $H_6$. Pour tester l'existence de *r* contraintes linéaires indépendantes sur les composantes du vecteur $\underset{(k,1)}{\beta}$, de la forme $H_0$ : $\underset{(r,k)}{C} . \underset{(k,1)}{\beta} = \underset{(r,1)}{q}$, on peut utiliser le résultat suivant :

[9] $r = (n-k+r) - (n-k) = dl_c - dl_{nc}$

**Si H₀ est vraie** alors : $f = \frac{\frac{SCR_c - SCR_{nc}}{dl_c - dl_{nc}}}{\frac{SCR_{nc}}{dl_{nc}}} \to F_{(dl_c - dl_{nc}, dl_{nc})}$ (21)

où $SCR_c$ et $SCR_{nc}$ représentent respectivement les sommes des carrés des résidus des estimations contrainte et non contrainte, $dl_c$ et $dl_{nc}$ étant les degrés de liberté de ces deux estimations.

### Écriture de la statistique du test en fonction des coefficients de détermination

Le coefficient de détermination peut être écrit en fonction de la somme des carrés des résidus estimés, compte tenu de l'expression (6). Dans le cas présent, puisque la variance totale est la même dans les deux modèles, nous avons :

$R_c^2 = 1 - \frac{SCR_c}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$ et $R_{nc}^2 = 1 - \frac{SCR_{nc}}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$

**Écriture de la statistique *f* en fonction des coefficients de détermination**


On en déduit une autre façon de calculer la statistique du test et donc la propriété suivante :

**Propriété** : Soit un modèle linéaire multiple de la forme $\underset{(n,1)}{y} = \underset{(n,k)}{X} . \underset{(k,1)}{\beta} + \underset{(n,1)}{\varepsilon}$, muni des hypothèses H₂, H₃ et H₆. Pour tester l'existence de *r* contraintes linéaires indépendantes sur les composantes du vecteur $\underset{(k,1)}{\beta}$, de la forme H₀ : $\underset{(r,k)}{C} . \underset{(k,1)}{\beta} = \underset{(r,1)}{q}$, on peut utiliser le résultat suivant :

**Si H₀ est vraie** alors : $f = \frac{(R_{nc}^2 - R_c^2)}{1 - R_{nc}^2} . \frac{n-k}{r} \to F_{(r, n-k)}$ (22)

où $R_c^2$ et $R_{nc}^2$ représentent respectivement les coefficients de détermination des estimations contrainte et non contrainte.

### Mise en œuvre du test : résumé des différentes étapes

Concrètement, la procédure à adopter pour mener le test de *r* contraintes linéaires sur les composantes du vecteur $\underset{(k,1)}{\beta}$ est la suivante :

1. Bien identifier le nombre *r* de contraintes linéairement indépendantes ;
2. Procéder aux estimations contrainte et non contrainte du modèle ;
3. Calculer la valeur de la statistique *f* à l'aide des expressions (21) ou (22) ;
4. Choisir une valeur pour le risque de première espèce $\alpha$ ;
5. Chercher dans une table de la loi de Fisher, la valeur critique $f_\alpha^*$, définie par :

$$Prob\{F_{(r,n-k)} > f_\alpha^*\} = \alpha$$

6. Prendre une <u>décision</u> : si $f > f_\alpha^*$ on rejette l'hypothèse nulle, c'est-à-dire qu'on considère que $\underset{(r,k)}{C} . \underset{(k,1)}{\beta} \neq \underset{(r,1)}{q}$. Dans le cas contraire, on ne rejette pas $H_0$ et les contraintes sont donc pertinentes.

## Cas particulier : le test de Fisher

Le test de Fisher fait partie des sorties automatiques fournies par les logiciels quand on effectue une estimation. C'est un cas particulier du test de *r* contraintes linéaires sur les composantes du vecteur $\underset{(k,1)}{\beta}$.

Dans le cadre d'un modèle de la forme :

$$y_i = \beta_0 + \beta_1 . x_{1,i} + \beta_2 . x_{2,i} + \cdots + \beta_{k-1} . x_{k-1,i} + \varepsilon_i$$

<u>L'hypothèse testée est la suivante</u> :

$$H_0 : [\beta_1 = \beta_2 = \cdots = \beta_{k-1} = 0]$$

On se demande si, à part la constante, tous les coefficients peuvent être considérés comme nuls. Dans le cas où on retiendrait l'hypothèse nulle, cela signifierait qu'aucune des variables explicatives figurant dans la spécification du modèle n'explique le phénomène *y*. On parle pour cette raison de **test de significativité globale de la régression**.

Plus que le coefficient de détermination, la mise en œuvre de ce test permet de juger de la qualité d'une régression.

### Définition du test

Au début de cette section, nous avons vu que l'hypothèse $H_0$ correspond à $(k-1)$ contraintes linéairement indépendantes qui peuvent être mises sous la forme :

$$\begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} . \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \\ \vdots \\ \vdots \\ \beta_{k-1} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

avec $\underset{(k-1,k)}{C} = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$ et $\underset{(k-1,1)}{q} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$.

<u>On est ici dans le cas où</u> : $r = k - 1$.

<u>Pour tester l'hypothèse</u> $\underset{(k-1,k)}{C} . \underset{(k,1)}{\beta} = \underset{(k-1,1)}{q}$, nous allons utiliser l'expression de *f* fournie par (22), soit :

$f = \frac{(R_{nc}^2 - R_c^2)}{1-R_{nc}^2} . \frac{n-k}{r}$ avec $R_c^2 = 0$. En effet, le modèle contraint s'écrit dans le cas envisagé ici : $y_i = \beta_0 + \varepsilon_i$. Comme nous l'avons vu précédemment l'estimation de ce modèle conduira à une valeur prédite pour la variable dépendante constante :

$$\hat{y}_i = \hat{\beta}_0 = \bar{y} \qquad \forall i$$

La variance expliquée par ce modèle est donc nulle et $R_c^2 = 0$.

Le test de significativité globale de la régression s'effectue alors en appliquant la règle (22) à notre cas particulier, soit :

**Si $H_0$ est vraie** alors : $f = \frac{R_{nc}^2}{1-R_{nc}^2} . \frac{n-k}{k-1} \rightarrow F_{(k-1,n-k)}$ (23)

Comme le *R*², calculé automatiquement à l'issue de chaque régression, la statistique permettant de mener le test de Fisher est aussi calculée systématiquement.

#### Remarque

Les sorties informatiques fournies par les logiciels d'économétrie permettent d'effectuer le test de Fisher à la simple lecture du listing.

La statistique *f* figurant dans l'expression (23) est toujours calculée, ainsi que le risque de première espèce $\alpha_f$ associé à la valeur de *f* (la *p-value* du test). $\alpha_f$, est appelé selon les logiciels « Probability of *f* » ou « Prob > *F* » et défini par :

$\alpha_f = Prob\{F_{(k-1,n-k)} > f\}$ où *f* est calculée à partir de l'expression figurant dans (23).

Pour un risque de première espèce $\alpha$, on rejettera l'hypothèse nulle, c'est-à-dire qu'on acceptera l'idée qu'au moins une variable explicative influence *y*, dès que $\alpha_f < \alpha$.

### Exemple

Considérons l'équation de salaire déjà étudiée. On a $T = 159$ et $k = 9$. Le logiciel utilisé (Eviews) fournit la valeur $f = 196{,}767$ (on peut vérifier que $f = \frac{R^2}{1-R^2} . \frac{150}{8}$ avec $R^2 = 0{,}913$).

Pour une loi de Fisher $F_{(8,150)}$, on trouve dans une table statistique la valeur critique associée à un risque de première espèce égal à 5% : $f^*_{0,05} = 2{,}00$. Puisque $196{,}8 \gg 2{,}00$, $f \gg f^*_{0,05}$ , et l'hypothèse nulle est donc rejetée. La croissance des salaires est expliquée par au moins une des variables explicatives envisagées dans la spécification estimée.

De plus, on remarque que le risque de première espèce associé à la valeur de *f* est égal à zéro, aux arrondis près. On trouve dans le listing $\alpha_f = \text{Prob} > F = 0{,}000$. Comme nous l'avons déjà souligné, la **donnée de la *p-value* $\alpha_f$, suffit pour réaliser le test de Fisher de façon quasi instantanée**. $\alpha_f$, ici égal à 0, est inférieur à n'importe quelle valeur $\alpha$ que l'on peut se fixer pour le risque de première espèce. La régression est donc significative.

## Exemple de test de contraintes linéaires sur l'équation de salaire

<u>La spécification estimée était la suivante</u> :

$$\dot{w}_t = b_1.tc_{t-1} + b_2.\dot{pc}_t + b_3.\dot{pc}_{t-1} + b_4.\dot{pc}_{t-2} + b_5.\dot{pc}_{t-3} + b_6.(wpq)_{t-1} + b_7.D821 + b_8.D823 + b_9 + u_t$$

Nous avons déjà analysé les résultats de l'estimation de cette équation et le test de Fisher a montré que la régression était globalement significative. D'autre part, les tests de significativité des coefficients ont conclu que l'on pouvait considérer les coefficients $b_4$, $b_5$, $b_7$ et $b_8$ comme non significativement différents de zéro.

À ce point de la réflexion, on peut se demander si la croissance des salaires nominaux est parfaitement indexée sur la croissance des prix à la consommation. Toutes choses égales par ailleurs, notamment à taux de chômage donné, cela signifierait que le pouvoir d'achat des salaires est maintenu, malgré l'inflation. Puisque $b_4 = b_5 = 0$, cette idée se traduit par la contrainte $b_2 + b_3 = 1$.

Comme on l'a vu, intégrer des contraintes dans un modèle permet d'améliorer la précision de l'estimation. Mais il convient de tester la pertinence des contraintes envisagées pour ne pas générer de biais.

<u>Dans l'exemple étudié ici, l'hypothèse testée est la suivante</u> :

$$\text{H}_0 : [b_2 + b_3 = 1 \text{ et } b_4 = b_5 = b_7 = 0]^{10}$$

On constate que H0 correspond à $r = 4$ contraintes linéaires indépendantes sur les coefficients de l'équation de salaire. On peut en effet écrire H0 sous la forme $\underset{(r,k)}{C} . \underset{(k,1)}{b} = \underset{(r,1)}{q}$ de la façon suivante :

$$\begin{pmatrix} 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix} . \begin{pmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \\ b_5 \\ b_6 \\ b_7 \\ b_8 \\ b_9 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

avec $\underset{(4,9)}{C} = \begin{pmatrix} 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}$ et $\underset{(4,1)}{q} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$

Pour tester H0, on estime le modèle contraint, puis on calcule la statistique *f* à l'aide de l'une des expressions (21) ou (22).

Pour obtenir l'estimateur contraint $\hat{b}_c$, on estime le modèle suivant :

$$\dot{w_t} = b_1 . tc_{t-1} + b_2 . \dot{pc}_t + b_3 . \dot{pc}_{t-1} + b_6 . (wpq)_{t-1} + b_8 . D823 + b_9 + w_t$$

en imposant la contrainte $b_2 + b_3 = 1$.

Autrement dit, on impose « manuellement » les contraintes $b_4 = b_5 = b_7 = 0$ en supprimant les variables correspondantes de la liste des variables explicatives, mais, pour la contrainte $b_2 + b_3 = 1$, on demande au logiciel de procéder à l'estimation sous contrainte selon la procédure décrite au début de cette section.

---

[10] On aurait pu inclure $b_8 = 0$ dans cette hypothèse nulle.

Les résultats de l'estimation du modèle contraint sont donnés ci-dessous :

Estimation d'une équation de salaire sous contrainte

Variable expliquée : $\dot{w}_t$

| | |
|---|---|
| $T = 159$ | $dl = 154$ |
| $R^2 = 0,911$ | $\overline{R}^2 = 0,908$ |
| $SCR = 0,0102$ | $\sigma = 0,0081$ |
| $f = 394,084$ | $Prob > F = 0,000$ |

| Variable | Coefficient estimé | Écart-type estimé | Valeur du $t$ | Prob > $\lvert t \rvert$ |
|---|---|---|---|---|
| $tc_{t-1}$ | $-0,1009$ | $0,0197$ | $-5,1160$ | $0,000$ |
| $\dot{pc}_t$ | $0,7432$ | $0,0755$ | $9,8448$ | $0,000$ |
| $\dot{pc}_{t-1}$ | $0,2568$ | $0,0755$ | $3,4010$ | $0,001$ |
| $(wpq)_{t-1}$ | $-0,0580$ | $0,0148$ | $-3,9095$ | $0,000$ |
| $D823$ | $-0,0051$ | $0,0082$ | $-0,6253$ | $0,533$ |
| Constante | $-0,3194$ | $0,0845$ | $-3,7801$ | $0,000$ |

Données trimestrielles pour l'économie française.
Période : 1951-T2 à 1990-T4.

$\dot{w}_t$ : taux de croissance du salaire nominal.
$tc_{t-1}$ : taux de chômage du trimestre précédent.
$\dot{pc}_t$ : taux de croissance des prix à la consommation du trimestre $t$.
$\dot{pc}_{t-1}$ : taux de croissance des prix à la consommation du trimestre $t-1$.
$(wpq)_{t-1}$ : terme de rappel.
$D823$ : variable indicatrice du troisième trimestre de 1982.

Modèle estimé :

$$\dot{w}_t = b_1 tc_{t-1} + b_2 \dot{pc}_t + b_3 \dot{pc}_{t-1} + b_6 (wpq)_{t-1} + b_8 D823 + +b_9 + w_t$$

sous la contrainte $b_2 + b_3 = 1$.

*Figure 2 : estimation d'une équation de salaire sous contrainte*

On peut effectuer les commentaires suivants :

- Comme prévu, la contrainte est exactement vérifiée au niveau des estimations obtenues : $\hat{b}_2 + \hat{b}_3 = 0{,}7432 + 0{,}2568 = 1$ ;

- Bien que le modèle contraint contienne 6 paramètres, le nombre de degrés de liberté de la régression contrainte, $dl_c$, est égal à 154, à cause de la contrainte $b_2 + b_3 = 1$ liant les paramètres $b_2$ et $b_3$ ;
- La régression contrainte, ayant un nombre moins élevé de variables explicatives, conduit bien à une diminution du $R^2$ et à une augmentation de la somme des carrés des résidus estimés. On a $R_c^2 = 0{,}911$ et $SCR_c = 0{,}0102$ alors que pour le modèle non contraint on avait $R_{nc}^2 = 0{,}913$ et $SCR_{nc} = 0{,}0099$ ;
- L'imposition des contraintes améliore bien la précision de l'estimation : les écart-types estimés des coefficients estimés sont inférieurs dans le modèle contraint par rapport à ceux de l'estimation non contrainte ;

| Écart-types estimés | Modèle contraint | Modèle non contraint |
|---|---|---|
| $\hat{\sigma}(\hat{b}_1)$ | 0,0197 | 0,0201 |
| $\hat{\sigma}(\hat{b}_2)$ | 0,0755 | 0,0837 |
| $\hat{\sigma}(\hat{b}_3)$ | 0,0755 | 0,0993 |
| $\hat{\sigma}(\hat{b}_6)$ | 0,0148 | 0,0169 |
| $\hat{\sigma}(\hat{b}_8)$ | 0,0082 | 0,0082 |
| $\hat{\sigma}(\hat{b}_9)$ | 0,0845 | 0,0967 |

*Tableau 2 : écart types - modèles contraint et non contraint*

- Pour tester la pertinence de l'ensemble des contraintes regroupées dans $H_0$, on calcule *f* à l'aide d'une des expressions (21) ou (22). On obtient, avec l'expression (21) :

$$f = \frac{\dfrac{SCR_c - SCR_{nc}}{dl_c - dl_{nc}}}{\dfrac{SCR_{nc}}{dl_{nc}}} = \frac{\dfrac{0{,}0102 - 0{,}0099}{154 - 150}}{\dfrac{0{,}0099}{150}} = 1{,}136$$

On sait que si $H_0$ est vraie, la statistique $f \rightarrow F_{(4,150)}$. Soit $f_{0,05}^*$ la valeur critique de cette loi de Fisher telle que :

$$Prob\{F_{(4,150)} > f_{0,05}^*\} = 0{,}05$$

On trouve, dans une table de la loi de Fisher la valeur : $f_{0,05}^* = 2{,}432$.

Puisque $f = 1{,}136$, $f < f_{0,05}^*$ et donc on ne rejette pas l'hypothèse nulle $H_0$.

**Décision** : La croissance des salaires nominaux est bien totalement indexée sur l'inflation (avec une indexation répartie sur 2 trimestres) et la non significativité des coefficients $b_4$, $b_5$ et $b_7$ est confirmée.

# Le test de Chow[11] : présentation et extensions

Appelé aussi **test de changement structurel**, le test de Chow permet d'examiner si les coefficients d'une régression sont stables par rapport aux observations utilisées. Plus exactement, on compare les estimations effectuées sur deux (ou plus de deux) sous-ensembles d'observations. Le test examine si les différences entre les coefficients estimés sont significatives.

Sur des séries chronologiques, les sous-ensembles d'observations correspondent à un découpage en périodes de l'échantillon initial. On parle dans ce cas de **test de stabilité temporelle** de la régression.

Sur des données en coupe transversale, on peut comparer des résultats obtenus sur des pays, des départements, des secteurs industriels différents. Concernant les individus, on peut s'intéresser à des résultats par classes d'âge, par sexe, par niveaux d'éducation. Dans ce cas, le test de Chow est souvent qualifié de **test d'homogénéité des comportements**.

Ce test est une simple application du test de *r* contraintes linéaires sur les paramètres. Il ne présente donc aucune nouveauté méthodologique par rapport à ce que nous venons d'étudier. Par contre, l'application des tests de Chow pose quelques problèmes particuliers dont il faut connaître la résolution.

## Méthode du test de Chow

Soit le modèle linéaire multiple $\underset{(n,1)}{y} = \underset{(n,k)}{X} . \underset{(k,1)}{\beta} + \underset{(n,1)}{\varepsilon}$. L'échantillon initial, qui comprend *n* observations, peut être partagé en deux sous-échantillons comprenant respectivement $n_1$ et $n_2$ observations.

Dans le cas d'un test de stabilité temporelle, les sous-échantillons correspondent à un découpage en périodes.

[11] Gregory C. Chow, « Tests of Equality Between Sets of Coefficients in Two Linear Regressions », *Econometrica*, vol. 28(3), 1960, p. 591–605

Pour un test d'homogénéité des comportements, ils correspondent à des secteurs, des pays, des catégories d'individus distincts.

Sur la première période (ou le premier secteur), le modèle s'écrit :

$$\underset{(n_1,1)}{y_1} = \underset{(n_1,k)}{X_1} \cdot \underset{(k,1)}{\beta_1} + \underset{(n_1,1)}{\varepsilon_1} \text{ avec } E[\varepsilon_1\varepsilon_1'] = \sigma^2 I_{n_1}$$

Sur la seconde période (ou le second secteur), le modèle s'écrit :

$$\underset{(n_2,1)}{y_2} = \underset{(n_2,k)}{X_2} \cdot \underset{(k,1)}{\beta_2} + \underset{(n_2,1)}{\varepsilon_2} \text{ avec } E[\varepsilon_2\varepsilon_2'] = \sigma^2 I_{n_2}$$

On a : $n = n_1 + n_2$, $n_1 > k$ et $n_2 > k$.

De plus : $\underset{(n,1)}{y} = \begin{pmatrix} \underset{(n_1,1)}{y_1} \\ \underset{(n_2,1)}{y_2} \end{pmatrix}$ ; $\underset{(n,k)}{X} = \begin{pmatrix} \underset{(n_1,k)}{X_1} \\ \underset{(n_2,k)}{X_2} \end{pmatrix}$ ; $\underset{(n,1)}{\varepsilon} = \begin{pmatrix} \underset{(n_1,1)}{\varepsilon_1} \\ \underset{(n_2,1)}{\varepsilon_2} \end{pmatrix}$

Soulignons que les hypothèses $H_2$, $H_3$ et $H_6$ sont maintenues. L'hypothèse $H_6$ implique en particulier que $E[\varepsilon\varepsilon'] = \sigma^2 I_n$ et donc que $E\left[\underset{(n_1,1)}{\varepsilon_1} \underset{(1,n_2)}{\varepsilon_2'}\right] = \underset{(n_1,n_2)}{0}$[12].

Si les comportements sont différents dans les deux sous-périodes (ou secteurs) on doit donc avoir : $\underset{(k,1)}{\beta_1} \neq \underset{(k,1)}{\beta_2}$. Si en revanche les comportements sont stables (ou homogènes) tout au long de la période (ou pour la totalité de l'échantillon), on doit avoir : $\underset{(k,1)}{\beta_1} = \underset{(k,1)}{\beta_2} = \underset{(k,1)}{\beta}$.

Le test de stabilité temporelle de la régression (ou d'homogénéité des comportements) se présente donc de la façon suivante :

$$\begin{cases} H_0 : \quad \underset{(n,1)}{y} = \underset{(n,k)}{X} \cdot \underset{(k,1)}{\beta} + \underset{(n,1)}{\varepsilon} \\ contre\ \bar{H} \ : \begin{cases} \underset{(n_1,1)}{y_1} = \underset{(n_1,k)}{X_1} \cdot \underset{(k,1)}{\beta_1} + \underset{(n_1,1)}{\varepsilon_1} \\ \underset{(n_2,1)}{y_2} = \underset{(n_2,k)}{X_2} \cdot \underset{(k,1)}{\beta_2} + \underset{(n_2,1)}{\varepsilon_2} \end{cases} \end{cases} \qquad (24)$$

Sous l'hypothèse alternative $\bar{H}$, on a deux modèles différents, un pour chaque sous-période (ou chaque secteur). Nous allons voir que l'écriture de ces deux modèles sous la forme d'un unique « sur-modèle » permet de ramener le problème à un simple test de *r* contraintes linéaires sur les paramètres. L'hypothèse alternative $\bar{H}$ peut s'écrire sous la forme :

[12] On suppose ici que les covariances entre les perturbations des deux sous-échantillons sont nulles.

$$\underset{(n,1)}{y} = \begin{pmatrix} \underset{(n_1,1)}{y_1} \\ \underset{(n_2,1)}{y_2} \end{pmatrix} = \begin{pmatrix} \underset{(n_1,k)}{X_1} & \underset{(n_2,k)}{0} \\ \underset{(n_1,k)}{0} & \underset{(n_2,k)}{X_2} \end{pmatrix} . \begin{pmatrix} \underset{(k,1)}{\beta_1} \\ \underset{(k,1)}{\beta_2} \end{pmatrix} + \begin{pmatrix} \underset{(n_1,1)}{\varepsilon_1} \\ \underset{(n_2,1)}{\varepsilon_2} \end{pmatrix}$$

soit :

$$\underset{(n,1)}{y} = \underset{(n,2k)}{\Psi} . \underset{(2k,1)}{B} + \underset{(n,1)}{\varepsilon} \qquad (25)$$

où $\underset{(n,2k)}{\Psi} = \begin{pmatrix} \underset{(n_1,k)}{X_1} & \underset{(n_2,k)}{0} \\ \underset{(n_1,k)}{0} & \underset{(n_2,k)}{X_2} \end{pmatrix}$ et $\underset{(2k,1)}{B} = \begin{pmatrix} \underset{(k,1)}{\beta_1} \\ \underset{(k,1)}{\beta_2} \end{pmatrix}$.

Le vecteur $\underset{(n,1)}{y}$ de la variable expliquée est le même sous l'hypothèse nulle de l'expression (24) et dans le sur-modèle (25).

En revanche, **les matrices des variables explicatives diffèrent dans l'une et l'autre hypothèse**. La matrice $\underset{(n,k)}{X}$ des variables explicatives du modèle sous l'hypothèse nulle correspond au simple empilement des matrices $\underset{(n_1,k)}{X_1}$ et $\underset{(n_2,k)}{X_2}$ , alors que la matrice $\underset{(n,2k)}{\Psi}$ des variables explicatives du sur-modèle est égale à la matrice bloc-diagonale formée à partir de ces deux même matrices.

Dans l'hypothèse où les comportements diffèrent d'une période à l'autre (ou entre les secteurs), le modèle s'écrit sous la forme (25), soit $\underset{(n,1)}{y} = \underset{(n,2k)}{\Psi} . \underset{(2k,1)}{B} + \underset{(n,1)}{\varepsilon}$ avec $\underset{(2k,1)}{B} = \begin{pmatrix} \underset{(k,1)}{\beta_1} \\ \underset{(k,1)}{\beta_2} \end{pmatrix}$.

Dans ce cadre, l'hypothèse nulle de stabilité (ou d'homogénéité) des comportements ( $\underset{(k,1)}{\beta_1} = \underset{(k,1)}{\beta_2}$ ) peut s'écrire comme l'hypothèse de *k* contraintes linéaires sur les coefficients du vecteur $\underset{(2k,1)}{B}$ de la façon suivante :

$$\underset{(k,2k)}{C} . \underset{(2k,1)}{B} = \underset{(k,1)}{0} \quad \text{avec} \quad \underset{(k,2k)}{C} = [I_k \,;\, -I_k]$$

où $I_k$ est la matrice identité d'ordre *k*.

Tester $\underset{(k,1)}{\beta_1} = \underset{(k,1)}{\beta_2}$ revient à tester $\underset{(k,2k)}{C} . \underset{(2k,1)}{B} = \underset{(k,1)}{0}$ . Notre problème se ramène donc à un test de $r = k$ contraintes linéaires sur les coefficients du sur-modèle (25) :

$$\begin{cases} H_0 : & \begin{cases} \underset{(n,1)}{y} = \underset{(n,2k)}{\Psi} . \underset{(2k,1)}{B} + \underset{(n,1)}{\varepsilon} \\ \text{sc} \ \underset{(k,2k)}{C} . \underset{(2k,1)}{B} = \underset{(k,1)}{0} \end{cases} \\ contre\ \overline{H} : & \underset{(n,1)}{y} = \underset{(n,2k)}{\Psi} . \underset{(2k,1)}{B} + \underset{(n,1)}{\varepsilon} \end{cases} \qquad (26)$$

Pour effectuer ce test, on utilisera le résultat donné dans l'expression (21) et pour calculer la statistique du test, il suffit de connaître les valeurs des sommes des carrés des résidus estimés et des degrés de liberté des modèles contraint et non contraint.

Estimer le modèle contraint revient simplement à estimer le modèle initial $\underset{(n,1)}{y} = \underset{(n,k)}{X} \cdot \underset{(k,1)}{\beta} + \underset{(n,1)}{\varepsilon}$ sur toutes les observations. On déduit de cette estimation $SCR_c$ et $dl_c$.

<u>L'estimation du modèle non contraint est réalisée en utilisant la propriété triviale suivante</u> :

**<u>Propriété</u> :** Estimer le modèle non contraint (25) revient à estimer le modèle initial période par période (ou secteur par secteur).

On a alors $SCR_{nc} = SCR_1 + SCR_2$ où $SCR_1$ est la somme des carrés des résidus estimés issus de l'estimation du modèle sur la première période (ou le premier secteur) et $SCR_2$ la somme des carrés des résidus estimés issus de l'estimation du modèle sur la seconde période (ou le second secteur). De plus, on a $dl_{nc} = dl_1 + dl_2$ où $dl_1$ et $dl_2$ sont respectivement les degrés de liberté des modèles relatifs à la première et à la seconde période (ou secteur).

En utilisant cette propriété, le test de Chow s'applique très simplement.

Supposons que l'on souhaite tester la stabilité des coefficients entre deux périodes ou étudier les différences de comportements entre deux secteurs (notés 1 et 2), la procédure à adopter est la suivante :

1. Estimer le modèle période par période (ou secteur par secteur). En déduire les valeurs de $SCR_1$, $SCR_2$, $dl_1 = n_1 - k$ et $dl_2 = n_2 - k$ ;
2. Estimer le modèle sur la période entière (ou en empilant les observations relatives aux deux secteurs). En déduire les valeurs de $SCR_c$ et $dl_c = n - k$ ;
3. Calculer la valeur de la statistique *f* à l'aide de l'expression (21). On a :

$$f = \frac{\frac{SCR_c - SCR_{nc}}{dl_c - dl_{nc}}}{\frac{SCR_{nc}}{dl_{nc}}} = \frac{\frac{SCR_c - (SCR_1 + SCR_2)}{k}}{\frac{(SCR_1 + SCR_2)}{n - 2k}}$$

Le test est basé sur le résultat suivant :

**Si $H_0$ est vraie** alors : $f \rightarrow F_{(k,n-2k)}$ (27)

4. Choisir une valeur pour le risque de première espèce $\alpha$ ;
5. Chercher dans une table de la loi de Fisher, la valeur critique $f_\alpha^*$, définie par :

$$Prob\{F_{(k,n-2k)} > f_\alpha^*\} = \alpha$$

6. Prendre la <u>décision</u> : si $f > f_{\alpha}^{*}$ on rejette l'hypothèse nulle. Autrement dit, on considère que $\underset{(k,1)}{\beta_1} \neq \underset{(k,1)}{\beta_2}$ : les coefficients se sont modifiés significativement entre les deux périodes (ou les comportements diffèrent significativement dans les deux secteurs).

**Remarque**

Comme nous l'avons déjà signalé, la plupart des logiciels économétriques ont des commandes permettant d'effectuer automatiquement les tests de contraintes linéaires sur les coefficients. Dans la pratique il suffit donc d'estimer le modèle sous la forme du sur-modèle :

$\begin{pmatrix} \underset{(n_1,1)}{y_1} \\ \underset{(n_2,1)}{y_2} \end{pmatrix} = \begin{pmatrix} \underset{(n_1,k)}{X_1} & \underset{(n_2,k)}{0} \\ \underset{(n_1,k)}{0} & \underset{(n_2,k)}{X_2} \end{pmatrix} . \begin{pmatrix} \underset{(k,1)}{\beta_1} \\ \underset{(k,1)}{\beta_2} \end{pmatrix} + \begin{pmatrix} \underset{(n_1,1)}{\varepsilon_1} \\ \underset{(n_2,1)}{\varepsilon_2} \end{pmatrix}$ et d'écrire la commande permettant de tester $\underset{(k,1)}{\beta_1} = \underset{(k,1)}{\beta_2}$.

Le seul piège à éviter concerne la constante car elle est générée automatiquement par les logiciels. En l'absence d'une commande spécifique, la matrice des variables explicatives du sur-modèle estimé comportera donc aussi un vecteur $\underset{(n,1)}{e_n}$, dont toutes les composantes sont égales à 1. Ceci revient à supposer que la constante $\beta_0$ est identique sur les deux périodes (ou pour les deux secteurs), ce qui, nous le verrons plus tard, peut créer un biais. Il convient donc de créer pour $\underset{(n,2k)}{\Psi}$, des vecteurs permettant d'estimer des constantes propres à chaque période (ou à chaque secteur). La méthode consiste à définir les variables indicatrices suivantes :

$I_{1i} = \begin{cases} 1 \text{ si } i \text{ appartient au groupe } 1 \\ 0 \text{ si } i \text{ appartient au groupe } 2 \end{cases}$ et $I_{2i} = \begin{cases} 1 \text{ si } i \text{ appartient au groupe } 2 \\ 0 \text{ si } i \text{ appartient au groupe } 1 \end{cases}$

Pour effectuer le test de Chow, il faudra alors commander une régression <u>sans constante</u> de $\underset{(n,1)}{y}$ sur $\underset{(n,2k)}{\Psi}$. Parmi les vecteurs de $\underset{(n,2k)}{\Psi}$ devront figurer les vecteurs $\underset{(n,1)}{I_1}$ et $\underset{(n,1)}{I_2}$ représentatifs des variables indicatrices.

## Hétéroscédasticité entre blocs

Dans ce qui précède, nous avons supposé que la variance de la perturbation était constante et égale à $\sigma^2$ sur les deux périodes (ou pour les deux secteurs).

Or, il est raisonnable d'admettre qu'en toute généralité, le test devrait se présenter de la façon suivante :

$$\begin{cases} H_0 : \underset{(n,1)}{y} = \underset{(n,k)}{X} . \underset{(k,1)}{\beta} + \underset{(n,1)}{\varepsilon} & \text{avec } E[\varepsilon\varepsilon'] = \sigma^2 I_n \\ contre\ \overline{H} : \begin{cases} \underset{(n_1,1)}{y_1} = \underset{(n_1,k)}{X_1} . \underset{(k,1)}{\beta_1} + \underset{(n_1,1)}{\varepsilon_1} & \text{avec } E[\varepsilon_1\varepsilon_1'] = \sigma_1^2 I_{n_1} \\ \underset{(n_2,1)}{y_2} = \underset{(n_2,k)}{X_2} . \underset{(k,1)}{\beta_2} + \underset{(n_2,1)}{\varepsilon_2} & \text{avec } E[\varepsilon_2\varepsilon_2'] = \sigma_2^2 I_{n_2} \end{cases} \end{cases}$$

Comme précédemment, on a : $n = n_1 + n_2$, $n_1 > k$, $n_2 > k$ et on suppose que $E\left[\underset{(n_1,1)(1,n_2)}{\varepsilon_1 \quad \varepsilon_2'}\right] = \underset{(n_1,n_2)}{0}$.

Pour ramener le test de stabilité des coefficients à un test de *k* contraintes linéaires sur les paramètres du sur-modèle (25), nous avons dû supposer : $\sigma_1^2 = \sigma_2^2 = \sigma^2$. En effet, la construction de la statistique *f* qui permet d'effectuer le test d'existence de contraintes linéaires sur les coefficients repose sur l'hypothèse que la perturbation a la même variance $\sigma^2$ dans les modèles contraint et non contraint.

Or, si on envisage la possibilité d'une rupture au premier ordre ($\beta_1 \neq \beta_2$), il est logique d'admettre l'éventualité d'une rupture au deuxième ordre de la forme $\sigma_1^2 \neq \sigma_2^2$, cas qualifié d'hétéroscédasticité entre blocs. On est généralement confronté à ce cas de figure sur des données en coupe transversale, quand on procède par exemple à des comparaisons internationales ou intersectorielles.

Divers travaux ont établi que, lorsque $\sigma_1^2 \neq \sigma_2^2$, le test de Chow décrit précédemment est biaisé dans le sens d'une surestimation du seuil de significativité du test. Autrement dit, on aura tendance à rejeter trop souvent l'hypothèse nulle $\beta_1 = \beta_2$. De ce fait, si l'application du test de Chow conduit à conclure que $\beta_1 = \beta_2$, une telle conclusion est robuste par rapport à l'existence ou non d'une hétéroscédasticité entre blocs des perturbations. En revanche, si on conclut que $\beta_1 \neq \beta_2$, il est important de tester l'hétéroscédasticité des perturbations.

## Test d'hétéroscédasticité des perturbations

Pour tester : $H_0 : \sigma_1^2 = \sigma_2^2$, on utilise une procédure correspondant au test de Goldfeld et Quandt[13]. On considère les hypothèses $H_2$, $H_3$ et une hypothèse $H_{6bis}$ qui correspond à un simple aménagement de l'hypothèse $H_6$ pour tenir compte de l'hétéroscédasticité potentielle des perturbations :

. $H_{6bis}$ : $\underset{(n_1,1)}{\varepsilon_1} \to N(0, \sigma_1^2 . I_{n_1})$ ; $\underset{(n_2,1)}{\varepsilon_2} \to N(0, \sigma_2^2 . I_{n_2})$ et $E\left[\underset{(n_1,1)(1,n_2)}{\varepsilon_1 \quad \varepsilon_2'}\right] = \underset{(n_1,n_2)}{0}$.

[13] S.M. Goldfeld et R.E. Quandt (1965), "Some Tests for Homoscedasticity". *Journal of the American Statistical Association* 60, 539–547

On suppose ainsi qu'il y a homoscédasticité à l'intérieur de chaque sous-échantillon, mais hétéroscédasticité entre les blocs. De plus, les perturbations sont supposées non corrélées entre elles.

Sous les hypothèses $H_2$, $H_3$ et $H_{6bis}$ on sait que :

$$\frac{\hat{\varepsilon}_1'\hat{\varepsilon}_1}{\sigma_1^2} \to \chi^2_{n_1-k} \qquad \text{et} \qquad \frac{\hat{\varepsilon}_2'\hat{\varepsilon}_2}{\sigma_2^2} \to \chi^2_{n_2-k}$$

où $\hat{\varepsilon}_1$ et $\hat{\varepsilon}_2$ sont les résidus estimés des régressions effectuées sur l'une et l'autre périodes (ou l'un et l'autre secteurs). Par ailleurs, $\underset{(n_1,1)}{\varepsilon_1}$ et $\underset{(n_2,1)}{\varepsilon_2}$ sont indépendants puisqu'il s'agit de vecteurs gaussiens de covariance nulle. On en déduit que $\underset{(n_1,1)}{\hat{\varepsilon}_1}$ et $\underset{(n_2,1)}{\hat{\varepsilon}_2}$, qui sont fonctions de $\varepsilon_1$ et $\varepsilon_2$, sont aussi indépendants. Les deux statistiques précédentes constituent donc des variables de Khi-deux indépendantes.

<u>On en déduit</u> :

$$\frac{\dfrac{\hat{\varepsilon}_1'\hat{\varepsilon}_1}{\sigma_1^2}\cdot\dfrac{1}{n_1-k}}{\dfrac{\hat{\varepsilon}_2'\hat{\varepsilon}_2}{\sigma_2^2}\cdot\dfrac{1}{n_2-k}} \to F_{(n_1-k,n_2-k)}$$

Sous $H_0$, on a $\sigma_1^2 = \sigma_2^2 = \sigma^2$ et le résultat précédent devient :

$$\varphi_1 = \frac{\dfrac{\hat{\varepsilon}_1'\hat{\varepsilon}_1}{n_1-k}}{\dfrac{\hat{\varepsilon}_2'\hat{\varepsilon}_2}{n_2-k}} \to F_{(n_1-k,n_2-k)}$$

Sous $H_0$, on constate aisément que l'on a aussi :

$$\varphi_2 = \frac{\dfrac{\hat{\varepsilon}_2'\hat{\varepsilon}_2}{n_2-k}}{\dfrac{\hat{\varepsilon}_1'\hat{\varepsilon}_1}{n_1-k}} \to F_{(n_2-k,n_1-k)}$$

On reconnaît aux numérateurs et aux dénominateurs des statistiques $\varphi_1$ et $\varphi_2$, les expressions des estimateurs $\hat{\sigma}_1^2$ et $\hat{\sigma}_2^2$ des variances $\sigma_1^2$ et $\sigma_2^2$ de la perturbation dans les deux groupes.

Finalement, le test d'hétéroscédasticité des perturbations entre les deux blocs (périodes, pays, secteurs, etc.) est basé sur le résultat suivant :

**Si $H_0$ est vraie** alors : $\varphi_1 = \frac{\hat{\sigma}_1^2}{\hat{\sigma}_2^2} \to F_{(n_1-k,n_2-k)}$ et $\varphi_2 = \frac{\hat{\sigma}_2^2}{\hat{\sigma}_1^2} \to F_{(n_2-k,n_1-k)}$ (28)

Soient $f^*_{\alpha,1}$ et $f^*_{\alpha,2}$ les valeurs critiques associées aux deux lois de Fisher figurant dans l'expression (28), en toute généralité et pour un risque de première espèce $\alpha$, on rejettera l'hypothèse nulle si :

$$\varphi_1 > f^*_{\alpha,1} \text{ ou } \varphi_2 > f^*_{\alpha,2}$$

Notons que le risque de première espèce associé à un tel test est égal à $2\alpha$.

En pratique, on procède à un test unilatéral de la forme :

$$\begin{array}{r} H_0 : \sigma_1^2 = \sigma_2^2 \\ contre\ \overline{H} :\ \sigma_1^2 > \sigma_2^2 \end{array} \qquad \text{ou} \qquad \begin{array}{r} H_0 : \sigma_2^2 = \sigma_1^2 \\ contre\ \overline{H} :\ \sigma_2^2 > \sigma_1^2 \end{array}$$

En effet, on a toujours une des deux variances estimées $\hat{\sigma}_1^2$ ou $\hat{\sigma}_2^2$ supérieure à l'autre. Une des deux statistiques $\varphi_1$ ou $\varphi_2$ est donc nécessairement inférieure à l'unité. Elle est alors obligatoirement inférieure à sa valeur critique et ne peut donc pas contribuer au rejet de l'hypothèse nulle[14].

**En résumé, la démarche à adopter est donc la suivante** :

1. Estimer le modèle sur les deux blocs d'observations et en déduire les valeurs de $\hat{\sigma}_1^2$ et de $\hat{\sigma}_2^2$ ;
2. Comparer les variances estimées des perturbations. Si $\hat{\sigma}_1^2 > \hat{\sigma}_2^2$, on procède au test unilatéral suivant[15] :

$$\begin{array}{r} H_0 : \sigma_1^2 = \sigma_2^2 \\ contre\ \overline{H} :\ \sigma_1^2 > \sigma_2^2 \end{array}$$

3. Calculer $\varphi_1 = \frac{\hat{\sigma}_1^2}{\hat{\sigma}_2^2}$. Le test est basé sur le résultat (28) :

**Si H0 est vraie** alors : $\varphi_1 = \frac{\hat{\sigma}_1^2}{\hat{\sigma}_2^2} \rightarrow F_{(n_1-k,n_2-k)}$

4. Choisir une valeur pour le risque de première espèce $\alpha$ ;
5. Chercher dans une table de la loi de Fisher la valeur critique $f^*_{\alpha,1}$, définie par :

$$Prob\{F_{(n_1-k,n_2-k)} > f^*_{\alpha,1}\} = \alpha$$

---

14 La plus petite valeur critique d'une loi de Fisher est obtenue lorsque les deux degrés de liberté sont infinis et cette valeur est égale à 1.

15 Bien entendu, si on observait $\hat{\sigma}_2^2 > \hat{\sigma}_1^2$, on procèderait au test unilatéral :

$$\begin{array}{r} H_0 : \sigma_2^2 = \sigma_1^2 \\ contre\ \overline{H} :\ \sigma_2^2 > \sigma_1^2 \end{array}$$

6. Prendre la <u>décision</u> : si $\varphi_1 > f^*_{\alpha,1}$ on rejette l'hypothèse nulle. Autrement dit, on considère que l'on est dans un cas d'hétéroscédasticité entre blocs.

### Test de Chow dans le cas d'hétéroscédasticité entre blocs

Lorsque le test de Goldfeld et Quandt conduit à rejeter l'hypothèse nulle (et donc à considérer que $\sigma_1^2 \neq \sigma_2^2$), il convient de définir une procédure permettant de mener le test de Chow dans ce cas.

Supposons que nous connaissions la valeur du rapport : $\omega = \frac{\sigma_1}{\sigma_2}$.

Dans ce cas, une simple transformation du modèle permet de construire un test de Chow non biaisé. Il suffit de multiplier le modèle relatif au second groupe d'observations par $\omega$ pour se retrouver dans un cas d'homoscédasticité globale. En effet, le modèle transformé s'écrit :

$$\omega . \underset{(n_2,1)}{y_2} = \omega . \underset{(n_2,k)}{X_2} . \underset{(k,1)}{\beta_2} + \omega . \underset{(n_2,1)}{\varepsilon_2} \qquad (29)$$

Posons $\underset{(n_2,1)}{v_2} = \omega . \underset{(n_2,1)}{\varepsilon_2}$ et calculons la nouvelle matrice de variances-covariances des perturbations.

$$\underset{(n_2,n_2)}{\mathrm{Var}}\ v_2 = E[v_2 v_2'] = E[\omega\varepsilon_2(\omega\varepsilon_2)'] = E[\omega^2\varepsilon_2\varepsilon_2'] = \omega^2 E[\varepsilon_2\varepsilon_2'] = \left(\frac{\sigma_1}{\sigma_2}\right)^2 \sigma_2^2 I_{n_2} = \sigma_1^2 I_{n_2}$$

Ainsi, la transformation (29) conduit, sur le deuxième bloc, à un modèle dont la variance des perturbations est constante et égale à $\sigma_1^2$, comme celle des perturbations du premier bloc.

Le test de Chow se mène alors en testant l'égalité $\underset{(k,1)}{\beta_1} = \underset{(k,1)}{\beta_2}$ sur le sur-modèle transformé suivant :

$$\begin{pmatrix} \underset{(n_1,1)}{y_1} \\ \omega . \underset{(n_2,1)}{y_2} \end{pmatrix} = \begin{pmatrix} \underset{(n_1,k)}{X_1} & \underset{(n_2,k)}{0} \\ \underset{(n_1,k)}{0} & \omega . \underset{(n_2,k)}{X_2} \end{pmatrix} . \begin{pmatrix} \underset{(k,1)}{\beta_1} \\ \underset{(k,1)}{\beta_2} \end{pmatrix} + \begin{pmatrix} \underset{(n_1,1)}{\varepsilon_1} \\ \omega . \underset{(n_2,1)}{\varepsilon_2} \end{pmatrix} \qquad (30)$$

En pratique, toutefois, on ne connaît pas la valeur exacte de $\omega$. On utilise alors une estimation $\widehat{\omega}$ de $\omega$, définie par : $\widehat{\omega} = \sqrt{\frac{\widehat{\sigma}_1^2}{\widehat{\sigma}_2^2}}$.

Le test de Chow est alors mené en testant l'égalité $\underset{(k,1)}{\beta_1} = \underset{(k,1)}{\beta_2}$ à partir du sur-modèle (30), en remplaçant $\omega$ par son estimation $\widehat{\omega}$.

<u>On procède ainsi en deux étapes</u> : la première étape consiste à estimer le modèle sur chacun des deux blocs afin d'obtenir les estimations $\widehat{\sigma}_1^2$ et $\widehat{\sigma}_2^2$. Dans la deuxième étape, on estime le sur-modèle transformé avec $\widehat{\omega}$.

Cette procédure revient à appliquer les moindres carrés quasi-généralisés.

Appliquer le test de Chow sur le sur-modèle transformé avec $\hat{\omega}$ n'est en principe correct qu'asymptotiquement, c'est-à-dire lorsque les deux sous-échantillons, correspondant aux deux blocs d'observations sont grands ($n_1 \to +\infty$ et $n_2 \to +\infty$). Dans la pratique, on utilise quand même cette méthode quand les deux sous-échantillons ne vérifient pas ces conditions.

### Retour sur la procédure du test de Chow

Puisqu'il faut tenir compte de l'éventualité d'une hétéroscédasticité entre blocs, la procédure à suivre pour mener le test de Chow est finalement la suivante :

1. Estimer le modèle sur les deux blocs d'observations et en déduire les valeurs de $\hat{\sigma}_1^2$ et de $\hat{\sigma}_2^2$ ;
2. Tester l'homoscédasticité entre blocs par la méthode de Goldfeld et Quandt ;
3. Si le test conduit à conclure que $\sigma_1^2 = \sigma_2^2$, appliquer simplement le test de Chow sur le sur-modèle (25) ;
4. Si le test conduit à conclure que $\sigma_1^2 \neq \sigma_2^2$, calculer $\hat{\omega}$ et appliquer le test de Chow sur le sur-modèle transformé (30), en remplaçant $\omega$ par son estimation $\hat{\omega}$.

## Un exemple de biais d'hétérogénéité

Comme nous l'avons remarqué au cours de la section précédente, imposer des contraintes non justifiées sur les coefficients conduit à biaiser les estimations. C'est le cas lorsqu'on postule une homogénéité des coefficients sur un ensemble d'observations correspondant à des populations hétérogènes. Le test de Chow vise à se prémunir contre ce type de biais, qualifié de **biais d'hétérogénéité**. Grâce à ce test, on examine s'il existe dans l'échantillon initial au moins deux sous-échantillons aux comportements distincts.

Un exemple intéressant de biais d'hétérogénéité est fourni par une étude[16] visant à estimer l'élasticité de substitution capital-travail (notée s) sur des coupes instantanées relatives à différents pays. Les auteurs prennent comme point de départ une fonction de production CES à élasticité de substitution constante de la forme :

$$Q_i = \left[\mu L_i^{-\rho} + (1-\mu) K_i^{-\rho}\right]^{-\frac{1}{\rho}}$$

[16] K.J. Arrow, H.B. Chenery, B.S. Minhas, and R.M. Solow, (1961), Capital-labor substitution and economic efficiency. Review of Economics and Statistics (43), p. 225-250.

Avec cette fonction de production, on sait que l'élasticité de substitution est égale à : $s = \frac{1}{1+\rho}$.

Pour obtenir la spécification estimée, les auteurs adoptent une hypothèse de concurrence pure et parfaite et supposent que l'entreprise représentative du secteur étudié pour le pays *i* maximise son profit. À l'équilibre, le logarithme de la production par tête est alors une fonction linéaire du logarithme du salaire réel, soit :

$$\text{Log}\left(\frac{Q}{L}\right)_i = s.\text{Log}\left(\frac{w}{p}\right)_i + \text{constante} + u_i \qquad (31)$$

Les auteurs effectuent 24 estimations de l'équation (30). Chaque estimation est réalisée sur une coupe instantanée comprenant 19 observations. Chacune de ces coupes correspond à l'observation, pour un secteur d'activité donné (sidérurgie par exemple), de la productivité du travail $\left(\frac{Q}{L}\right)_i$ et du salaire réel $\left(\frac{w}{p}\right)_i$ dans 19 pays $(i = 1, \dots, 19)$. Chacune des 24 estimations correspond à un des 24 secteurs considérés.

Un des enjeux de ce travail empirique était de déterminer si l'hypothèse d'une fonction de production Cobb-Douglas était une approximation correcte de la réalité. On sait que dans le cas d'une telle fonction de production, l'élasticité de substitution est égale à l'unité. Il suffisait donc de tester l'hypothèse $H_0$ : $[s = 1]$ à partir des estimations réalisées. Sur les 24 estimations mises en œuvre par les auteurs, 23 ont conduit à la conclusion que *s* était significativement inférieure à 1, impliquant un rejet de l'hypothèse Cobb-Douglas.

Ce travail a été ultérieurement critiqué dans la littérature économique. En reprenant les mêmes estimations, il s'avère que le résultat obtenu par K.J. Arrow et al. est affecté par un biais d'hétérogénéité, dû à la présence de deux catégories de pays dans les observations : les pays industrialisés et les pays en voie de développement. La prise en compte d'une valeur de la constante significativement différente entre ces deux catégories de pays permet d'aboutir à une estimation sans biais de la pente (s). En effet, on trouve alors que $s = 1$ pour les deux groupes de pays dans la plupart des secteurs (22 secteurs sur les 24 étudiés).

Le graphique suivant permet de comprendre comment le biais d'hétérogénéité peut affecter substantiellement l'estimation de l'élasticité de substitution *s*.

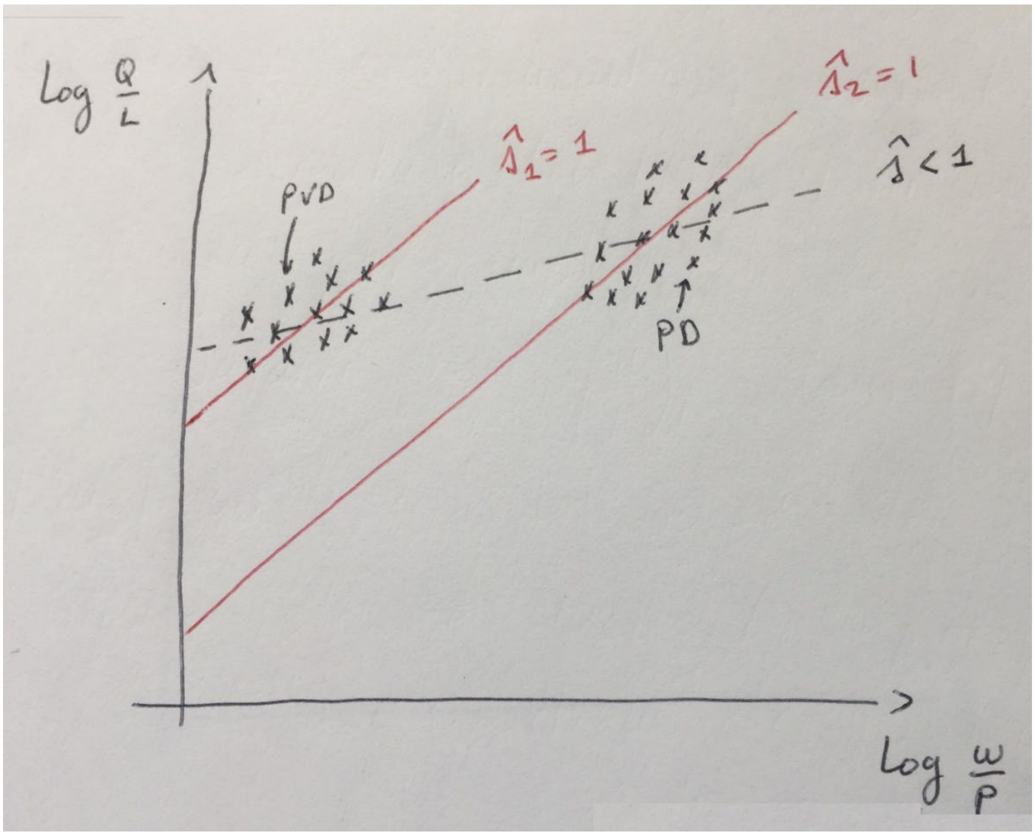

*Figure 3 : Estimation relative à un secteur d'activité*

Les pays développés (PD) ont des salaires réels et une productivité du travail plus élevés que les pays en voie de développement (PVD). En contraignant les paramètres à être égaux dans les deux types de pays, on trouve une droite de régression figurant en pointillés correspondant à $\hat{s} < 1$. En effectuant l'estimation sans contraindre les coefficients à être égaux dans les deux types de pays, on trouve des constantes estimées très différentes et des pentes estimées $\hat{s}_1$et $\hat{s}_2$. Les tests permettent d'établir que $s_1 = s_2 = 1$.

## Remarque concernant la mise en œuvre du test de Chow

L'exemple précédent permet d'effectuer une remarque pratique d'une grande importance pour la mise en œuvre du test de Chow.

Supposons que nous voulions tester l'égalité des pentes *s* entre les pays industrialisés et les pays en voie de développement. Dans ce cas, il faut se garder de contraindre les constantes à être égales dans les deux catégories de pays. Sinon, on se retrouve dans la situation illustrée ci-dessous, où la contrainte imposée sur les constantes, biaise les pentes estimées. Celles-ci semblent significativement différentes alors qu'en réalité elles sont égales :

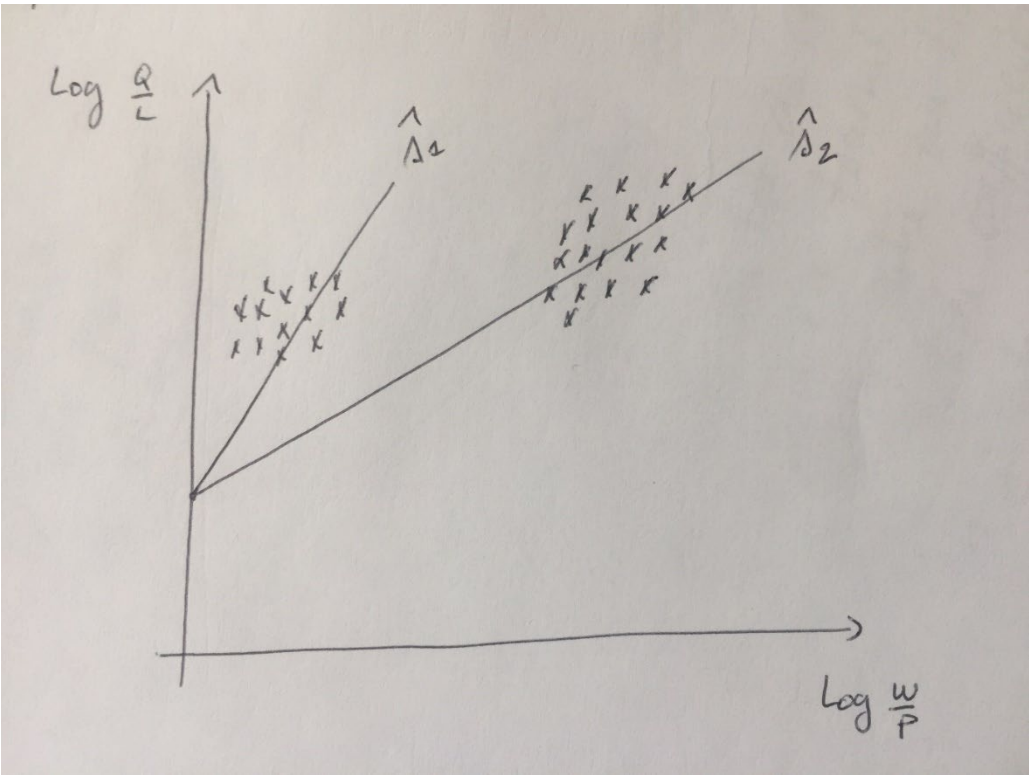


*Figure 4 : exemple de contrainte imposée sur les constantes qui biaise les pentes estimées*

<u>En pratique</u>, quand on veut tester l'homogénéité de certains coefficients, la démarche pertinente consiste à tester différentes combinaisons possibles de contraintes, **en allant du moins contraignant au plus contraignant**.

Considérons le modèle linéaire multiple avec $k = 3$ suivant :

$$\underset{(n,1)}{y} = \underset{(n,1)}{x} \cdot \underset{(1,1)}{a} + \underset{(n,1)}{z} \cdot \underset{(1,1)}{b} + \underset{(n,1)}{c} + \underset{(n,1)}{\varepsilon}$$

Ce modèle comporte effectivement 3 paramètres à estimer : *a*, *b* et $\underset{(n,1)}{c}$. Dans le cas où les observations comporteraient deux sous-ensembles aux comportements intégralement hétérogènes, on aurait :

$$\begin{cases} \underset{(n_1,1)}{y_1} = \underset{(n_1,1)}{x_1} \cdot \underset{(1,1)}{a_1} + \underset{(n_1,1)}{z_1} \cdot \underset{(1,1)}{b_1} + \underset{(n_1,1)}{c_1} + \underset{(n_1,1)}{\varepsilon_1} \\ \underset{(n_2,1)}{y_2} = \underset{(n_2,1)}{x_2} \cdot \underset{(1,1)}{a_2} + \underset{(n_2,1)}{z_2} \cdot \underset{(1,1)}{b_2} + \underset{(n_2,1)}{c_2} + \underset{(n_2,1)}{\varepsilon_2} \end{cases}$$

L'application de la démarche recommandée précédemment doit conduire à pratiquer successivement 7 tests de Chow, correspondant aux hypothèses nulles suivantes :

H01 : $[a_1 = a_2]$ H04 : $[a_1 = a_2 \text{ et } b_1 = b_2]$ H07 : $[a_1 = a_2 \text{ et } b_1 = b_2 \text{ et } c_1 = c_2]$

H02 : $[b_1 = b_2]$ H05 : $[a_1 = a_2 \text{ et } c_1 = c_2]$

H03 : $[c_1 = c_2]$ H06 : $[b_1 = b_2 \text{ et } c_1 = c_2]$

# Conclusion

L'ensemble des tests usuels sur le modèle linéaire multiple viennent d'être présentés de façon exhaustive. Ces tests ne remettent pas en cause les hypothèses sous-jacentes initialement posées. La remise en question de certaines de ces hypothèses et les conséquences sur les résultats obtenus par l'estimation seront l'objet des approfondissements proposés lors du troisième chapitre de ce cours.

# Références

### Comment citer ce cours ?

Introduction à l'économétrie, Olivier Baron, AUNEGe (http://aunege.fr), CC – BY NC ND (http://creativecommons.org/licenses/by-nc-nd/4.0/).



**Table des illustrations**

**Figures**



**Tableaux**