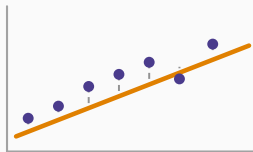


Introduction à l'économétrie

Chapitre 15 : L'hétéroscédasticité



Jaouad Madkour

Faculté d'Économie et de Gestion de Tanger · 27 août 2026

Voir le problème

Ce que cela coûte

La pratique actuelle

Ce qu'il faut retenir

Voir le problème

Homoscédasticité et son contraire

$$V(\varepsilon_i) = \sigma_\varepsilon^2 \text{ pour tout } i$$

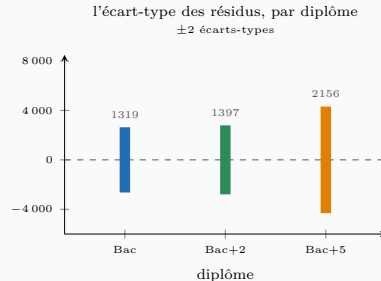
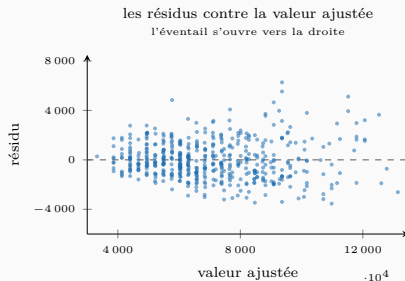
Quand cette variance dépend des variables explicatives, on parle d'**hétéroscédasticité** — littéralement, une dispersion inégale.

Pourquoi c'est la règle sur données individuelles

Les ménages pauvres consomment presque tout leur revenu : peu de latitude, donc peu de dispersion.

Les ménages aisés arbitrent entre consommation, épargne et placement : beaucoup de latitude, donc beaucoup de dispersion.

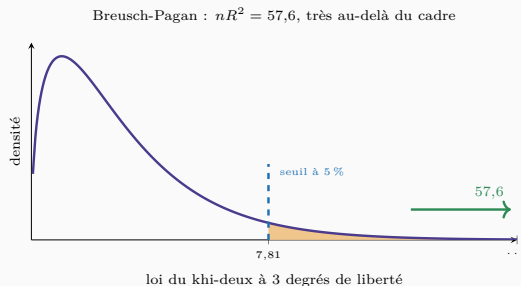
Le comportement économique lui-même produit l'hétéroscédasticité. Ce n'est pas un défaut des données.



La dispersion des salaires n'est pas la même en haut et en bas de l'échelle.

Chez les titulaires d'un Bac, les résidus s'écartent de 1 319 dirhams ; chez les Bac+5, de 2 156. L'hypothèse H3 est visiblement fausse — et cela n'a rien d'anormal : **sur des données individuelles, l'hétéroscédasticité est la règle, non l'exception.**

Les coefficients restent sans biais. Ce sont les *écarts-types* qui sont faux, donc tous les *t*, toutes les *p*-values et tous les intervalles. Le remède tient en une option de la commande d'estimation.



Le principe des deux tests

On régresse le **carré des résidus** sur les variables explicatives. Si elles l'expliquent, la variance dépend d'elles : c'est l'hétéroscédasticité.

Breusch-Pagan : sur les variables du modèle.

White : sur les variables, leurs carrés et leurs produits croisés — plus général, plus gourmand en degrés de liberté.

Statistique : nR^2 de cette seconde régression, à lire dans un khi-deux.

Sur nos données : 57,6 pour Breusch-Pagan, 48,6 pour White. Les deux rejettent massivement.

Un test qui rejette ne dit pas *combien* c'est grave : regardez toujours le graphique des résidus.

Les deux, sur le modèle M2

Test	Statistique	Degrés de liberté	Seuil
Breusch-Pagan	57,60	3	7,81
White	48,55	2	5,99

Le rejet est massif, et ce n'est pas une surprise

L'écart-type des résidus passe de **1 319** dirhams chez les titulaires d'un Bac à **2 156** chez les Bac+5 : un rapport de 1,6.

Le graphique avait déjà tout dit. Le test ne fait que le confirmer avec un chiffre.

Ce que cela coûte

L'inventaire

1. Les coefficients restent **sans biais** — H2 n'est pas en cause.
2. Ils ne sont plus de **variance minimale** : les MCO cessent d'être BLUE.
3. Les **écarts-types calculés sont faux**, donc tous les t , p -values et intervalles.

La troisième est la seule qui compte en pratique

On peut vivre avec une efficacité dégradée. On ne peut pas vivre avec des tests faux.

Et le sens de l'erreur n'est **pas prévisible** : les écarts-types peuvent être sous-estimés — on croit à une significativité qu'on n'a pas — ou surestimés.

Les coefficients ne bougent pas d'un centime. Seuls les écarts-types changent.

	Coefficient	Écart-type		Statistique t	
		MCO	robuste	MCO	robuste
Constante	3 320,42	158,85	144,37	20,90	23,00
Ancienneté	270,27	11,72	13,00	23,06	20,78
Bac+2	1 099,57	158,59	138,43	6,93	7,94
Bac+5	3 336,99	184,25	226,28	18,11	14,75

C'est exactement ce que la théorie annonçait : l'hétéroscédasticité ne biaise rien, elle fausse la précision.

La correction de White ne modifie pas l'estimation : elle recalcule sa variance sans supposer que le bruit est le même partout.

La ligne Bac+5 est la plus instructive. C'est le groupe le plus dispersé, et son écart-type passe de 184 à 226 : les MCO le sous-estimaient de 23 %, donc surestimaient sa significativité. Ailleurs la correction joue en sens inverse.

Rien ne garantit le sens du changement — c'est bien pourquoi on ne peut pas s'en passer par prudence. Sur données individuelles, l'usage est aujourd'hui de publier **systematiquement** les écarts-types robustes.

Les écarts-types de White

On ne suppose plus que la variance est la même partout : on l'**estime observation par observation** à partir du carré des résidus.

$$\widehat{V}(\hat{\beta}) = (X'X)^{-1} \left(\sum_i e_i^2 x_i x_i' \right) (X'X)^{-1}$$

Ce qu'il ne change pas, et ce qu'il change

Les coefficients ne bougent pas d'un centime : on ne réestime rien.

Seule la **variance** est recalculée. C'est une correction de la mesure d'incertitude, pas de l'estimation — et c'est exactement ce que la théorie annonçait.

Sur nos données, il joue dans les deux sens

Variable	t MCO	t robuste
Ancienneté	23,06	20,78
Bac+2	6,93	7,94
Bac+5	18,11	14,75

La ligne Bac+5 est la plus instructive

C'est le groupe le plus dispersé, et son écart-type passe de 184 à 226 : les MCO le **sous-estimaient de vingt-trois pour cent**, donc surestimaient sa significativité.

Ailleurs la correction joue en sens inverse. **On ne peut donc pas s'en passer par prudence** — il n'y a pas de sens « prudent ».

La pratique actuelle

L'usage professionnel

Sur données individuelles, on publie **systématiquement** des écarts-types robustes, sans tester au préalable.

La correction ne coûte rien, ne fait aucun mal si les erreurs sont homoscédastiques, et évite une faute grave si elles ne le sont pas.

Pourquoi apprendre les tests, alors

Parce qu'ils font comprendre le mécanisme, et parce qu'ils gardent leur utilité ailleurs : quand le remède coûte quelque chose, ou quand la forme de l'hétéroscédasticité intéresse en elle-même.

Un économiste qui constate que la dispersion des salaires croît avec le diplôme n'a pas trouvé un défaut : il a trouvé un fait économique.

L'hétéroscédasticité comme signal

Une hétéroscédasticité très marquée signale parfois une **forme fonctionnelle mal choisie**.

Sur nos salaires, passer au logarithme réduit sensiblement le problème : la dispersion des salaires croît avec le niveau, mais celle des *logarithmes* est bien plus stable. Le chapitre 19 y reviendra.

La leçon générale

Avant de corriger un symptôme, se demander s'il ne révèle pas une erreur de spécification.

Les écarts-types robustes réparent la conséquence ; le logarithme supprime parfois la cause.

Ce qu'il faut retenir

Six points

1. L'hétéroscédasticité est la **règle** sur données individuelles, non l'exception.
2. Sur nos salaires : écart-type des résidus de **1 319** à **2 156** dirhams selon le diplôme.
3. Breusch-Pagan **57,6** et White **48,6** : rejet massif, confirmé par le graphique.
4. Les coefficients restent **sans biais**; seuls les **écarts-types** sont faux.
5. Les écarts-types **robustes de White** corrigent, sans toucher aux coefficients.
6. Le sens de la correction n'est pas prévisible : la prudence ne dispense pas de corriger.

La suite

Nous changeons de jeu de données. Sur l'enquête de ménages, ce n'est plus la dispersion qui pose problème, c'est le **regroupement** : les observations n'y sont pas tombées une par une.