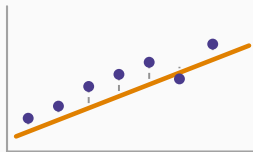


Introduction à l'économétrie

Chapitre 13 : La carte des violations



Jaouad Madkour

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

Une seule ligne compte vraiment

Le rôle des résidus

Comment les tests fonctionnent

Ce qu'il faut retenir

Une seule ligne compte vraiment

Une seule ligne est en orange, et c'est la seule qui compte vraiment.

Problème	Biais ?	Écarts-types	Comment le voir	Que faire
Variable omise H2 violée	OUI	faussés	réflexion économique aucun test ne la révèle	ajouter la variable ou instrumenter
Hétéroscédasticité H3 violée	non	faussés	résidus en éventail Breusch-Pagan, White	écarts-types robustes de White
Erreurs groupées H4 violée	non	faussés	résidus liés dans un groupe moyennes par groupe	écarts-types de grappe au niveau du regroupement
Multicolinéarité H5 presque violée	non	gonflés	VIF élevé coefficients instables	retirer, regrouper, ou ne rien faire
Non-normalité	non	justes si n grand	histogramme, Jarque-Bera	rien, si $n \geq 30$

Toutes les violations ne se valent pas.

Quatre d'entre elles ne touchent que les *écarts-types* : les coefficients restent justes en moyenne, et une correction technique suffit. On les détecte par un test, on les répare par une formule.

La variable omise, elle, biaise les coefficients eux-mêmes. Aucun test ne la signale — puisque la variable manquante n'est pas dans les données. Elle se détecte par le raisonnement économique, et se répare en changeant de modèle ou de méthode.

C'est pourquoi un travail économétrique sérieux consacre plus de temps à la *spécification* qu'à l'estimation.

Deux familles de problèmes

Ceux qui faussent les écarts-types : hétéroscédasticité, erreurs groupées, multicollinéarité. Les coefficients restent justes.

Celui qui fausse les coefficients : la variable omise. Rien d'autre ne compte autant.

La conséquence sur la façon de travailler

Les premiers se **détectent par un test** et se **réparent par une formule**. Ce sont des problèmes techniques, et ils ont des solutions techniques.

Le second se détecte par le **raisonnement économique** — car la variable manquante n'est pas dans les données — et se répare en changeant de modèle. **C'est un problème de spécification, et aucun logiciel ne le résoudra.**

Le rôle des résidus

Ce qu'on regarde

Les erreurs ε_i sont inconnues. Les résidus e_i sont calculés.

Tout diagnostic passe par les résidus : ils sont la seule trace observable de ce que le modèle n'explique pas.

Un principe qui ne varie pas

Sous les hypothèses, les résidus doivent ressembler à du **bruit pur** : aucune structure, aucune tendance, aucune relation avec quoi que ce soit.

Toute régularité visible dans les résidus est une information que le modèle a laissée échapper. C'est le fil conducteur des cinq chapitres qui suivent.

Les quatre graphiques à faire systématiquement

Avant tout commentaire des coefficients

Graphique	Ce qu'il révèle
Résidus contre valeurs ajustées	hétéroscédasticité, non-linéarité
Résidus moyens par groupe	erreurs groupées
Histogramme des résidus	non-normalité, valeurs extrêmes
Résidus contre chaque variable	forme fonctionnelle mal choisie

Ils coûtent quatre lignes de code

Aucun de ces quatre graphiques ne prend plus de dix secondes à produire, et l'un d'eux vaut souvent mieux qu'une batterie de tests.

Un test dit si un problème existe; un graphique dit à quoi il ressemble. Les deux sont nécessaires, et le graphique vient d'abord.

Comment les tests fonctionnent

La régression auxiliaire

La plupart des tests de diagnostic suivent le même schéma :

1. estimer le modèle et récupérer les résidus;
2. **régresser une fonction des résidus** sur quelque chose;
3. si cette seconde régression explique quelque chose, l'hypothèse est violée.

Trois tests, un seul principe

Test	On régresse	Sur
Breusch-Pagan	e_i^2	les variables explicatives
White	e_i^2	variables, carrés, produits
RESET	y_i	variables et puissances de $\hat{y}_i$

Toujours la même

$$nR_{\text{auxiliaire}}^2 \sim \chi_q^2$$

où q est le nombre de variables ajoutées dans la régression auxiliaire.

Pourquoi un khi-deux

Parce qu'on additionne des écarts au carré — c'est le signe reconnu au chapitre 19 de l'inférence.

Et le test est **unilatéral à droite** : seule une valeur élevée contredit l'hypothèse d'absence de problème.

Là encore, rien de nouveau : on applique un test appris ailleurs.

Ils sont sensibles à la taille de l'échantillon

Sur cinq cents observations, la moindre hétéroscédasticité devient significative. Sur trente, une hétéroscédasticité massive peut passer inaperçue.

Un test qui rejette ne dit pas que le problème est grave. Il dit qu'il existe. Le graphique dit s'il compte.

La bonne pratique

Sur données individuelles, l'usage actuel est de ne même plus tester l'hétéroscédasticité : on la suppose présente et l'on publie directement des écarts-types robustes.

Le test garde son intérêt pédagogique — il fait comprendre le mécanisme — et son intérêt pratique quand le remède coûte quelque chose.

Ce qu'il faut retenir

Six points

1. Quatre violations ne touchent que les **écarts-types**; une seule, la variable omise, biaise les **coefficients**.
2. Les premières se détectent par un **test** et se réparent par une **formule**.
3. La seconde se détecte par le **raisonnement économique** et exige de changer de modèle.
4. Tout diagnostic passe par les **résidus** : ils doivent ressembler à du bruit pur.
5. Quatre graphiques, systématiquement, avant tout commentaire de coefficient.
6. Les tests suivent tous le même schéma : une **régression auxiliaire** et un nR^2 lu dans un khi-deux.

La suite

La carte est dressée. Les cinq chapitres qui suivent la parcourent case par case, chacun sur le jeu de données où le problème se rencontre naturellement.

On commence par celui qui n'est pas vraiment une violation, et qui déroute le plus : la **multicolinéarité**.