

## Annexe B

Nous avons testé la stationnarité des différentes séries présentées dans le graphique 1. Les tests usuels (tests de stationnarité de Dickey-Fuller modifié et de Dickey-Fuller augmenté avec tendance, ou encore test de non-stationnarité de Kwiatkowski-Phillips-Schmidt-Shin) appliqués au montant des prescriptions par médecins en niveau et en différence première indiquent que cette variable est intégrée d'ordre 1. L'analyse de la série sur le nombre de visiteurs médicaux par médecins est rendue difficile du fait de son caractère lacunaire<sup>1</sup>. Nous avons mené les tests sur la variable originelle avec trous et sur une variable complétée par interpolation linéaire à partir des dates antérieures et postérieures aux dates où elle est manquante. Dans les deux cas, et toujours d'après les tests usuels<sup>2</sup>, la variable apparaît intégrée d'ordre deux. Le tableau 5 résume ces résultats, donne le degré d'intégration des autres variables explicatives et présente les notations utilisées par la suite.

---

<sup>1</sup> La série n'a été reconstruite que pour certaines années (graphique 1), celles où nous avons le plus de certitudes quant à la fiabilité de la source. Il n'y a donc pratiquement que pour la période 2002 à 2009 que nous disposons de plusieurs années consécutives permettant une analyse via un modèle autorégressif vectoriel avec inclusion de variables retardées.

<sup>2</sup> Résultats détaillés disponibles auprès des auteurs.

**Tableau 5. Notations et synthèse des tests de stationnarité des séries**

Série	Notation		Ordre d'intégration
	en niveau	en différence	
Montant des prescriptions par médecin libéral	$y_t$	$\Delta y_t$	1
Visiteurs médicaux par médecin	$x_t$	$\Delta x_t$	2
Visiteurs médicaux par médecin interpolé	$x'_t$	$\Delta x'_t$	2
Nombre d'actes par an et par médecin libéral	$z_{1t}$	$\Delta z_{1t}$	1
Indice de vieillissement de la population	$z_{2t}$	$\Delta z_{2t}$	2
Part des présentations de moins de cinq ans dans le chiffre d'affaires officinal	$z_{3t}$	$\Delta z_{3t}$	1

Note : Le tableau synthétise les résultats concordants des tests de Dickey Fuller modifié, Dickey-Fuller augmenté avec ou sans trend, et Kwiatkowski-Phillips-Schmidt-Shin appliqués aux séries, et à leurs différences premières (la variable moins sa valeur retardée) et seconde (la différence première moins la différence première retardée). Les résultats détaillés sont disponibles auprès des auteurs. Un niveau d'intégration d'ordre 1 signifie que la série est non stationnaire, mais que sa valeur retardée l'est. Un niveau d'intégration d'ordre 2 signifie que la différence première de la série n'est pas stationnaire, mais que la différence seconde l'est.

La différence d'ordre d'intégration entre les deux séries d'intérêt  $y_t$  et  $x_t$  rend théoriquement inapplicables les modèles usuels<sup>3</sup> pour ces deux seules variables, puisqu'une variable indépendante doit être au moins du même ordre d'intégration que la variable dépendante (Baffes, 1997 ; Banerjee *et al.*, 1993). La seule façon correcte de s'en sortir devrait être d'inclure comme variable explicative dans les modèles au moins une variable du même ordre que  $y_t$ , à savoir  $z_{1t}$  ou  $z_{3t}$  (tableau 5). Enfin, on considérera les variables en différence première de manière à rendre stationnaire au moins la variable dépendante  $\Delta y_t$  et à éviter les problèmes éventuels de cointégration<sup>4</sup>. On estime trois modèles VAR (« vecteurs autorégressifs ») incluant soit uniquement les séries  $\Delta x'_t$  et  $\Delta y_t$ , soit ces deux variables ainsi que  $\Delta z_{1t}$  (afin d'éviter les problèmes de non-stationnarité évoqués plus haut), soit l'ensemble des séries présentées dans le tableau 5 (afin de tester la robustesse des liens entre visites

<sup>3</sup> À la fois les modèles à correction d'erreur (utilisés si les séries sont cointégrées) et les modèles VAR (« vecteurs autorégressifs », dans le cas de séries non cointégrées).

<sup>4</sup> Les tests du multiplicateur de Lagrange appliqués aux modèles estimés en niveaux montrent que les résidus de ces modèles sont auto-corrélés, suggérant ainsi que ces modèles sont mal spécifiés.

médicales et prescriptions)<sup>5</sup>. Tous les modèles doivent être estimés avec la variable de visiteurs médicaux par médecins interpolée (sinon on ne disposerait que de six observations dans les années 2000). Les trois modèles incluent uniquement les deux premiers termes de retard des séries en différence. Le nombre optimal de retards peut être déterminé en examinant la valeur de l'erreur finale de prédiction (FPE), ou encore certains critères d'information existant dans la littérature – le plus connu par Akaike (AIC), mais aussi celui de Schwarz (SBIC) et de Hannan et Quinn (HQIC). L'analyse de l'erreur finale de prédiction suggère de considérer deux retards dans un modèle sans contrôle et un retard si des variables de contrôle du tableau 5 sont incluses dans l'analyse. L'analyse du critère AIC suggère d'utiliser deux retards dans un VAR sans contrôle ou incluant uniquement le nombre d'actes par an et par médecin comme variable additionnelle. Elle suggère d'autoriser uniquement un retard dans un modèle VAR incluant l'ensemble des variables du tableau 5. Enfin l'analyse des critères SBIC et HQIC suggère le plus souvent d'utiliser un retard. Au vu de ces différents éléments, nous présentons l'analyse de modèles VAR avec et sans contrôles, en autorisant systématiquement deux retards. Conformément aux préconisations issues de l'analyse des critères d'information statistique, nous avons également reproduit l'analyse avec un seul retard et vérifié que les résultats sont robustes et quantitativement similaires<sup>6</sup>.

Le modèle 1, même s'il est potentiellement mal spécifié du fait du caractère non stationnaire de  $\Delta x'_t$ , suggère un effet positif important des variations passées du nombre de visiteurs par médecin sur les variations contemporaines des volumes de prescription (tableau 6, modèle 1). Ce sont les variations retardées de deux ans qui semblent avoir l'effet le plus important : un visiteur médical de plus pour dix médecins libéraux est associé à une augmentation du volume de prescription par médecin d'environ 20 000 euros au bout d'un an et d'environ 50 000 euros par an au bout de deux ans (significatif au seuil de 1 %). L'analyse montre également que les variations annuelles des volumes de prescriptions sont auto-corrélées ( $\Delta y_{t-1}$  a un effet sur  $\Delta y_t$ ) et, de façon rassurante, qu'elles n'ont aucun effet sur l'intensité des visites. Ces constats sont robustes à l'inclusion d'autres séries, telles que le nombre d'actes par an et par médecin (tableau 6, modèle 2), ou encore l'indice de vieillissement de la population et la part des présentations de moins de cinq ans (tableau 6, modèle 3). Cette dernière variable a également dû être interpolée. Dans ce dernier modèle, l'effet des variations du nombre de visiteurs médicaux par médecins retardé de deux ans devient moins significatif d'un point de vue

---

<sup>5</sup> Les valeurs des facteurs d'inflation de la variance (*Variance Inflation Factors* ou VIF) obtenues pour les variables du modèle 3 sont très largement inférieures à 1, indiquant qu'il n'y a aucun problème de colinéarité entre ces variables, même dans le modèle le plus exhaustif.

<sup>6</sup> Résultats détaillés disponibles auprès des auteurs.

statistique, au détriment du premier retard qui devient presque significatif au seuil de 10 %. Quoiqu'il en soit, dans ce modèle et les précédents, les deux retards considérés ensemble expliquent très fortement les volumes de prescriptions (test de causalité de Granger). On peut noter également la stabilité des coefficients estimés pour le lien entre intensité de la visite et prescriptions sur l'ensemble des trois modèles testés<sup>7</sup>. Enfin, on n'observe pas d'autocorrélation des résidus dans ces modèles (tests du multiplicateur de Lagrange), suggérant qu'ils sont correctement spécifiés.

---

<sup>7</sup> Les mêmes modèles estimés en logarithme donnent également des résultats très cohérents dans leur ensemble et indiquent une élasticité de l'ordre de 0.45 à un an et 0.60 à deux ans des prescriptions au nombre de visiteurs médicaux par médecins : une hausse de 1 % du nombre de visiteurs médicaux par médecin est associée à une hausse de 0,45 % (resp. 0,60 %) des prescriptions par médecin un an (resp. deux ans) après.

**Tableau 6. Trois modèles VAR pour évaluer les liens entre visites médicales et prescriptions**

	Modèle 1		Modèle 2		Modèle 3	
	Coefficient	Écart-type	Coefficient	Écart-type	Coefficient	Écart-type
Effets sur les prescriptions par médecin :						
Effet de $\Delta y_{t-1}$ sur $\Delta y_t$	-0.20	0.17	-0.01	0.25	0.08	0.28
Effet de $\Delta y_{t-2}$ sur $\Delta y_t$	-0.11	0.17	-0.19	0.24	-0.29	0.26
Effet de $\Delta x_{t-1}$ sur $\Delta y_t$	215.25	199.72	228.23	197.60	297.14	205.43
Effet de $\Delta x_{t-2}$ sur $\Delta y_t$	586.49***	214.32	553.21***	213.85	456.58*	244.42
Effets sur le nombre de visiteurs médicaux par médecin :						
Effet de $\Delta y_{t-1}$ sur $\Delta x'_t$	0.00	0.00	0.00	0.00	0.00	0.00
Effet de $\Delta y_{t-2}$ sur $\Delta x'_t$	0.00	0.00	0.00	0.00	0.00	0.00
Effet de $\Delta x_{t-1}$ sur $\Delta x'_t$	0.49***	0.18	0.48***	0.17	0.49***	0.18
Effet de $\Delta x_{t-2}$ sur $\Delta x'_t$	0.20	0.19	0.19	0.18	0.10	0.21
Autocorrélation des résidus (test du multiplicateur de Lagrange)						
	Non (p=0.63)		Non (p=0.59)		Non (p=0.80)	
$\Delta x'_t$ cause $\Delta y_t$ au sens de Granger						
	Oui (p=0.001)		Oui (p=0.001)		Oui (p=0.008)	
Autres séries incluses						
	Non		Oui : $\Delta z_{1t}$		Oui : $\Delta z_{1t}, \Delta z_{2t}, \Delta z_{3t}$	
Nombre d'observations						
	37		37		37	

Notes : Le tableau présente tout (modèle 1) ou partie (modèles 2 et 3) de la matrice d'autocorrélation issue de l'estimation d'un modèle VAR avec deux retards. Les liens entre les séries « de contrôle »  $\Delta z_{1t}$ ,  $\Delta z_{2t}$  et  $\Delta z_{3t}$  ne sont pas présentés. \* : significatif au seuil de 10 %, \*\*\* : significatif au seuil de 1 %.

L'analyse par un modèle VAR nécessite d'utiliser la variable du nombre de visiteurs médicaux par médecins complétée par interpolation linéaire. La qualité d'une telle

interpolation n'étant pas garantie, on peut vouloir adopter une démarche plus pragmatique basée sur la seule exploitation de la variable non interpolée. Cette variable n'étant que rarement disponible à deux dates consécutives avant les années 2000, la seule estimation possible dans ce cas repose sur un modèle en niveau du type :

$$y_t = \alpha * y_{t-1} + \beta * x_t + \gamma_1 * z_{1t} + \gamma_2 * z_{2t} + \gamma_3 * z_{3t} + \epsilon_t \quad (*)$$

où un (ou plusieurs) retard de  $y_t$  peut éventuellement être inclus pour contrôler pour l'autocorrélation de cette seule variable. L'estimation de tels modèles sans inclusion de la variable dépendante retardée aboutit à des liens positifs significatifs entre  $y_t$  et  $x_t$ , quelles que soient les variables de contrôle  $z$  retenues. L'inclusion du terme retardé  $y_{t-1}$  (comme dans l'équation (\*)) rend les résultats un peu moins nets : on continue de trouver un lien positif entre  $y_t$  et  $x_t$  en l'absence de variables de contrôles et avec la plupart des combinaisons possibles de  $z_{1t}$ ,  $z_{2t}$  et  $z_{3t}$  comme contrôles. Cependant pour certaines combinaisons, et notamment lorsqu'on contrôle pour le nombre d'actes, la relation positive entre  $y_t$  et  $x_t$  disparaît. Ceci peut s'expliquer par la forte corrélation (0,94) entre le volume de prescription et le nombre d'actes qui induit des problèmes de colinéarité dans les modèles en niveau lorsque cette dernière variable est incluse. Les valeurs des facteurs d'inflation de la variance (VIF) pour les modèles qui incluent cette variable tendent à être supérieures à 10, voire à 20 (la variable  $z_{1t}$  a par exemple un VIF égal à 35 dans le modèle (\*)), suggérant que la colinéarité trop forte entre les variables rend difficile d'isoler l'effet de chacune d'entre elles. Ces problèmes de colinéarité plaident en faveur soit d'une analyse en niveau très parcimonieuse à partir de la variable non interpolée du nombre de visiteurs médicaux par médecins, soit d'une analyse en différence, à l'aide un modèle VAR approprié, à partir des séries prises en différence première. Dans les deux cas, on trouve un effet positif et largement significatif de l'intensité de la visite médicale sur le volume de prescription.