

Les allocations régionales de ressources en secteur ambulatoire : l'introduction des interactions entre médecine généraliste et médecine spécialisée

P. Crépon

CNAMTS - 66 avenue du Maine - 75014 Paris

L'allocation des ressources assurant le financement des soins de santé par le système de sécurité sociale pose un problème compliqué dès lors qu'à un niveau de découpage géographique donné, on souhaite harmoniser les enveloppes entre plusieurs secteurs différents de soins de santé intervenant dans le cadre du système sanitaire, de façon complémentaire ou concurrente.

La démarche qui est exposée ci-dessous entend apporter une réponse possible à cette préoccupation. Elle se situe, plus précisément dans le cadre suivant.

Les ressources concernées correspondent aux dépenses présentées au remboursement relatives à l'activité des omnipraticiens d'une part, des spécialistes d'autre part, en honoraires et en prescriptions.

Le niveau géographique où cette allocation est calculée est celui de la région administrative (22 régions pour la France métropolitaine).

1. Le cadre méthodologique de la démarche

La méthode de détermination des allocations régionales de ressources entre dans le groupe des méthodes dites explicatives, c'est-à-dire que le niveau de consommation unitaire de ces ressources sur chacun des deux champs - activité des omnipraticiens et activité des spécialistes -, sur la dernière année connue, est confronté à un certain nombre de facteurs explicatifs liés à la demande et à l'offre de soins, dits « **facteurs exogènes** », à travers un modèle.

1.1. Indépendamment des problèmes de transfert de consommation de ressources entre les deux champs susnommés, il existe sur chacun d'entre eux, un niveau théorique d'allocation de ressources.

Désignons par :

Y le niveau de consommation unitaire des ressources sur l'année N.

X1, X2, ...Xn les facteurs explicatifs exogènes de cette consommation

Nous avons une formulation générale de la forme :

$$Y = \alpha X1 + \beta X2 + \dots + \mu + \text{résidu}$$

$$\text{avec : } Y' = \alpha X1 + \beta X2 + \dots + \mu$$

et donc $Y = Y' + \text{résidu}$

Y' représente un niveau théorique d'allocation de ressources.

Le résidu représente l'écart qu'il convient de résorber pour passer de la situation actuelle à une situation « **cible** ».

1.2. La démarche précédente offre l'inconvénient de traiter de façon indépendante les allocations de ressources sur les deux champs, c'est-à-dire ne comporte pas de dispositif de régulation de la répartition des ressources entre ces deux champs, à partir du constat des interactions de consommation de ces ressources sur chacun d'entre eux.

Sous certaines conditions, le système suivant, dit d'équations simultanées, répond à cette objection, puisqu'il offre l'avantage d'intégrer automatiquement les transferts de consommation de ces ressources.

Reprenons les équations ci-dessus et affectons les aux deux champs, celui des omnipraticiens et celui des spécialistes.

$$Y_{\text{omni}} = Y'_{\text{omni}} + (\text{résidu})_1$$

$$Y_{\text{spe}} = Y'_{\text{spe}} + (\text{résidu})_1$$

Le système d'équations simultanées implique que les variables dépendantes théoriques, dites « **variables endogènes** », soient introduites comme facteurs explicatifs dans un deuxième système d'équations.

$$Y_{\text{omni}} = \varphi_1 \times Y'_{\text{spe}} + \gamma_1 \times X_1 + \eta_1 + (\text{résidu})_2$$

$$\text{avec } Y''_{\text{omni}} = \varphi_1 \times Y'_{\text{spe}} + \eta_1$$

$$Y_{\text{spe}} = \varphi_2 \times Y'_{\text{omni}} + \delta \times X_2 + \eta_2 + (\text{résidu})_2$$

$$\text{avec } Y''_{\text{spe}} = \varphi_2 \times Y'_{\text{omni}} + \delta \times X_2 + \eta_2$$

Les facteurs peuvent être ou non introduits et être identiques, mais de préférence différents des facteurs du premier niveau d'équations.

En pratique, aucun facteur exogène n'a été introduit au deuxième niveau d'équation

L'ensemble du système doit répondre à des **contraintes d'identification** lesquelles sont, au cas présent, respectées.

conditions nécessaires :

a - le système doit comprendre au plus **K x M** paramètres

K représentant le nombre de variables exogènes et **M** le nombre de variables endogènes.

Si le modèle est identifiable, il est alors juste identifiable.

b - la somme du nombre de paramètres liés aux endogènes M_i et du nombre de paramètres liés aux exogènes K_i doit être $<$ ou $=$ au nombre d'exogènes

$$M_1 \neq \varphi_1 = 1 ; K_1 \neq \gamma_1, \eta_1 = 2$$

Le niveau théorique acceptable d'allocation unitaire des ressources résulte ensuite d'une réduction des écarts entre Y_{omni} et Y''_{omni} , Y_{spe} et Y''_{spe}

Le cadre général de la démarche étant ainsi posé, revenons aux différentes composantes de cette problématique.

2. Les facteurs explicatifs de la consommation des ressources

2.1. Les facteurs explicatifs de la consommation liés à la demande

Nous avons privilégié deux types d'indicateurs :

2.1.1. Un indicateur de morbidité reposant sur le dénombrement théorique des personnes reconnues officiellement, au cours de l'année, comme atteintes d'une affection de longue durée. Il s'agit donc d'un flux et non pas d'un stock.

On passe de la notion de flux observé à celle de flux théorique à partir du raisonnement suivant :

La reconnaissance d'une affection de longue durée chez une personne déterminée doit être considérée comme l'indication d'une prise en charge de la personne, acceptée par la **Sécurité Sociale**, à partir :

- d'un diagnostic effectué par un médecin, spécialiste ou omnipraticien, du secteur libéral ou du secteur public,
- et de la confirmation de ce diagnostic par un médecin-conseil de la Sécurité Sociale.

La présence d'une affection de longue durée est, par ailleurs, largement conditionnée par l'âge de la personne.

Si bien qu'on peut passer de la notion d'ALD observées à celle d'ALD théoriques de la façon suivante :

$$ALD^{obs} = a \times \text{indic.âge} + b \times \text{indic.offre de soins} + \text{constante} + \text{résidu}$$

$$\text{avec } ALD^{obs} = ALD^{théor} + \text{résidu}$$

$$\text{donc } ALD^{théor} = a \times \text{indic.âge} + b \times \text{indic.offre de soins} + \text{constante}$$

Dans une situation de répartition optimale de l'offre de soins, la densité médicale qui constitue l'indicateur retenu pour l'offre de soins devrait donner à chaque personne le même droit de tirage dans des conditions d'âge équivalentes.

Donc dès lors que les coefficients **a** et **b** ont été déterminés, il suffit de remplacer dans l'équation précédente l'indicateur d'offre de soins par l'indicateur d'âge - sous réserve que dans la formulation l'ensemble des variables dépendantes ou indépendantes aient été centrées sur la valeur prise au plan national -.

Sous ces conditions,

$$ALD^{théor} = (a + b) \times \text{indic.âge} + \text{constante}$$

Les $ALD^{théor}$ constituent la variable X2 du premier système d'équation.

L'estimateur (a+b) est probablement biaisé car il ne rend pas compte des disparités de besoin d'accès aux soins à tranche d'âge équivalente.

Le dénombrement d'ALD théoriques, en tant que conduisant à la définition d'un indicateur de morbidité doit ensuite être confronté à un indicateur de mortalité. Nous avons retenu l'indicateur de mortalité brute, ce qui constitue, en quelque sorte une validation de cet indicateur.

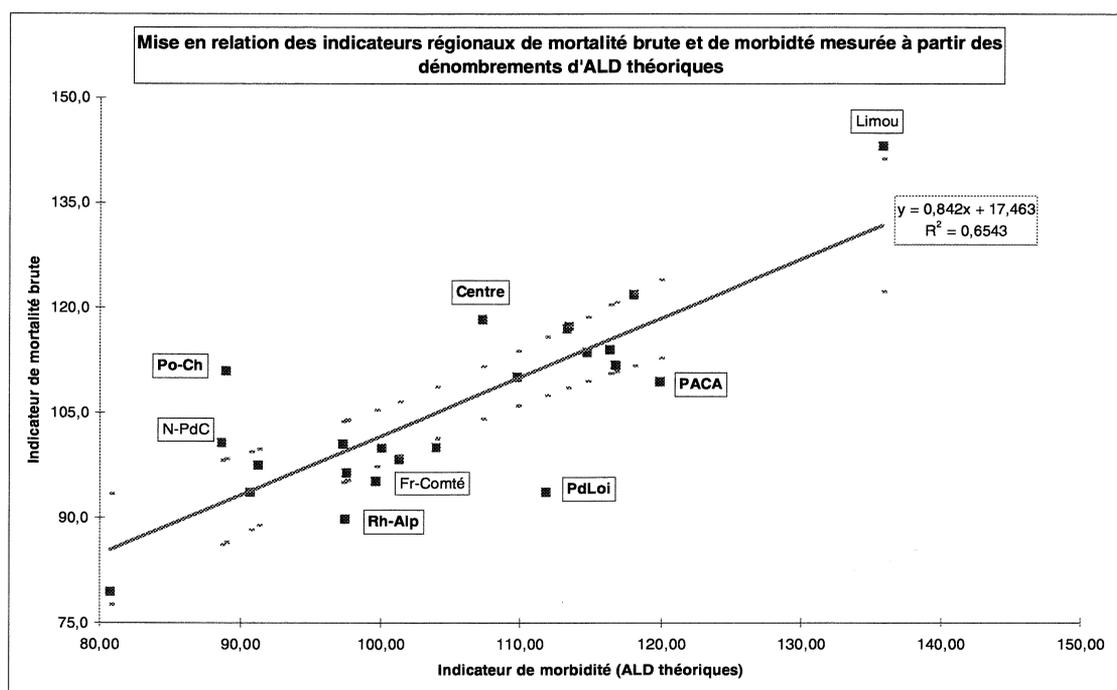
La régression linéaire : $\text{Indic. mortal.} = f(\text{indic. morbid.})$ fait apparaître :

un coefficient de détermination $R^2 = 0,66$, un t de Student = 3,5 ainsi qu'un coefficient directeur affecté à la variable $\text{indic. morbid} = 0,88$.

On peut donc en conclure que l'indicateur de morbidité établi à partir du dénombrement d'ALD constitue une proxi-variable satisfaisante de la mortalité, laquelle, comme indicateur, réunit les suffrages de la communauté scientifique.

Le graphique ci-dessous illustre, au niveau régional, le rapprochement entre morbidité et mortalité. Les pointillés rouges visualisent l'intervalle de confiance autour de la droite de tendance. Deux types de régions sortent significativement de cet intervalle :

- celles qui se situent au-dessus de l'intervalle de confiance ; soit la situation de morbidité y est inférieure à ce que laisserait supposer l'âge moyen de la population - Poitou-Charentes et Centre -, soit la morbidité y est beaucoup plus élevée que ce qui serait attendu, compte-tenu de l'âge de la population Nord-Pas-de-Calais,
- celles qui se situent au-dessous de l'intervalle de confiance, dont l'état de santé est sans doute caractérisé par des pathologies relativement moins létales que la moyenne - ou mieux soignées -.



La morbidité offre l'avantage de constituer un indicateur théorique dont on peut toujours réduire l'écart par rapport à un indicateur observé, donc de reporter, en partie, à ce niveau le problème de la réduction des écarts entre un niveau théorique d'allocation de ressources et le constat effectué dans le passé le plus récent du niveau de consommation de ces mêmes ressources.

Nous avons pris le parti de retenir, dans la mise en oeuvre des équations simultanées, un indicateur de morbidité s'appuyant sur une réduction de 5% de l'écart entre ALD observées et ALD théoriques.

2.1.2. Un indicateur de précarité qui prend en compte :

- d'une part, la fraction de la population bénéficiant de minima sociaux (allocations attribuées sous condition de ressources (RMI, API, AAH, FNS¹), **IP1^R**
- d'autre part, le revenu disponible par ménage, **IP2^R**

Si nous désignons par **IP1^N** et **IP2^N** ces mêmes indicateurs au niveau national un indicateur **IP^R** centré est fourni par la formule :

$$IP^R = (IP1^R / IP1^N) \times (IP2^N / IP2^R)$$

L'indicateur de précarité retenu ici gagnerait à être consolidé en affinant IP1 et IP2 grâce à une pondération par les populations concernées.

2.2. Les facteurs explicatifs de la consommation liés à l'offre

Le modèle se doit de prendre en compte les transferts de consommation en direction du secteur public.

Deux méthodes pouvaient être utilisées :

- soit considérer qu'un indicateur de transfert vers le secteur public constituait une variable explicative de la consommation des ressources en direction du secteur libéral dont la prise en compte s'effectuerait à l'identique des facteurs explicatifs liés à la demande,
- soit considérer que ces transferts de consommation permettaient de modifier la population concernée par la consommation de ressources en direction du secteur libéral, auquel cas cette prise en compte intervenait dans le calcul de la variable dépendante Y.

Eu égard au nombre limité d'observations (22), nous avons privilégié la seconde démarche afin de limiter le nombre de facteurs exogènes.

L'impact du secteur public a donc été apprécié par le rapport établi comme suit :

$$K = \text{Actes secteur libéral (C,V, K, Kc, Z)} / [\text{Actes secteur libéral} + \text{Consultations externes (secteur public)} + \text{Urgences (secteur public)}]$$

La population résidente de chaque région est multipliée par le coefficient K.

3. Les modalités alternatives de mise en application pratique de la démarche

L'allocation régionale des ressources, quels que soient le champ et la prise en compte ou non des transferts intersectoriels, conduit inévitablement à une confrontation entre un niveau observé de consommation de ces ressources et un niveau optimal d'allocation de celles-ci.

Mais, ce niveau optimal dépendant lui même d'un niveau optimal de satisfaction des besoins, donc de reconnaissance par le système sanitaire de la gravité de l'affection, il est possible de définir 2 modalités M1 et M2 :

M1 - soit se satisfaire **pour l'année à venir d'un niveau sous-optimal de satisfaction des besoins** ; ceci revient à calculer un indicateur de morbidité sur la base d'un dénombrement d'ALD, intermédiaire entre les ALD observées et les ALD théoriques, à partir, par exemple, d'une réduction

¹ RMI : revenu minimum d'insertion ; API : allocation pou parent isolé ; AAH : allocation pour adulte handicapé ; FNS : fond national de solidarité.

des écarts de 5% ; l'allocation théorique des ressources s'inscrit ensuite dans le cadre strict de l'indicateur de morbidité ainsi défini ; si l'écart entre l'allocation par personne des ressources et la consommation de ces ressources apparaît trop important, la réduction de cet écart sur une période pluriannuelle - qui n'est pas envisagée dans cette présentation - est toujours possible ; nous indiquerons dans les commentaires des résultats l'incidence d'une planification pluriannuelle ;

M2 - soit définir un niveau théorique d'allocation de ressources en fonction d'un niveau optimal de satisfaction des besoins et **ensuite seulement de se satisfaire d'un niveau sous-optimal d'allocation de ressources** en réduisant de 5% par exemple l'écart entre consommation des ressources et allocation théorique de celles-ci.

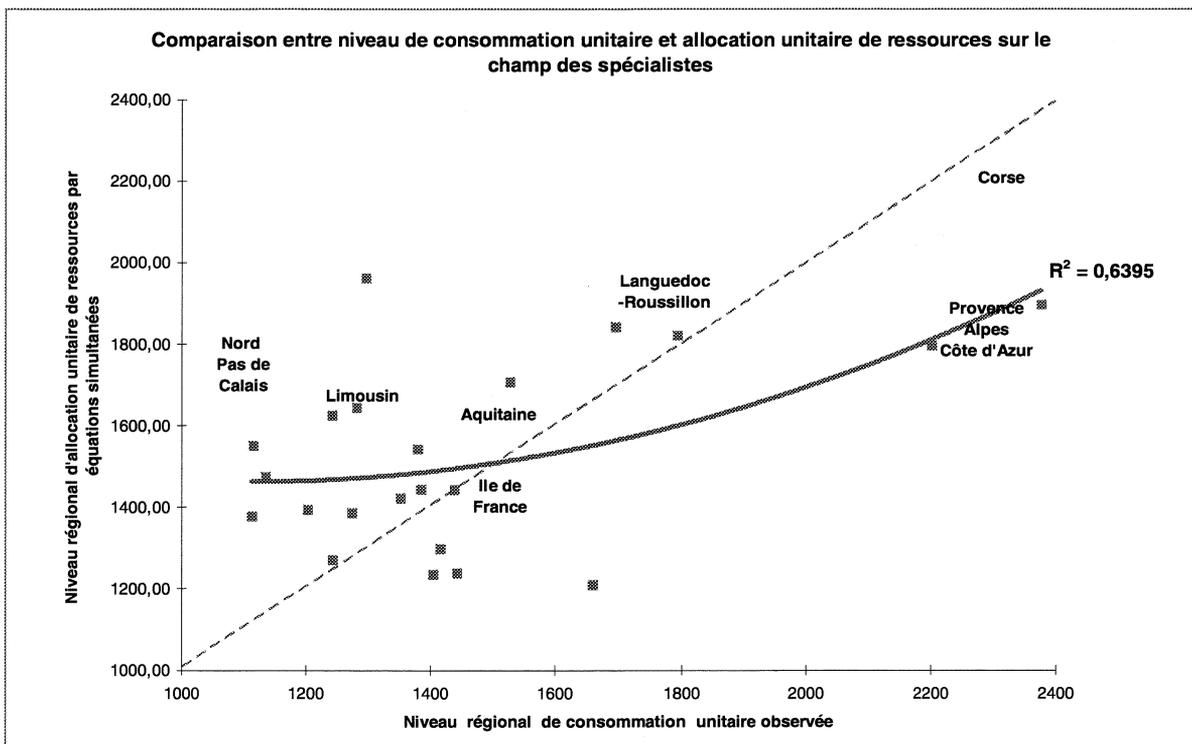
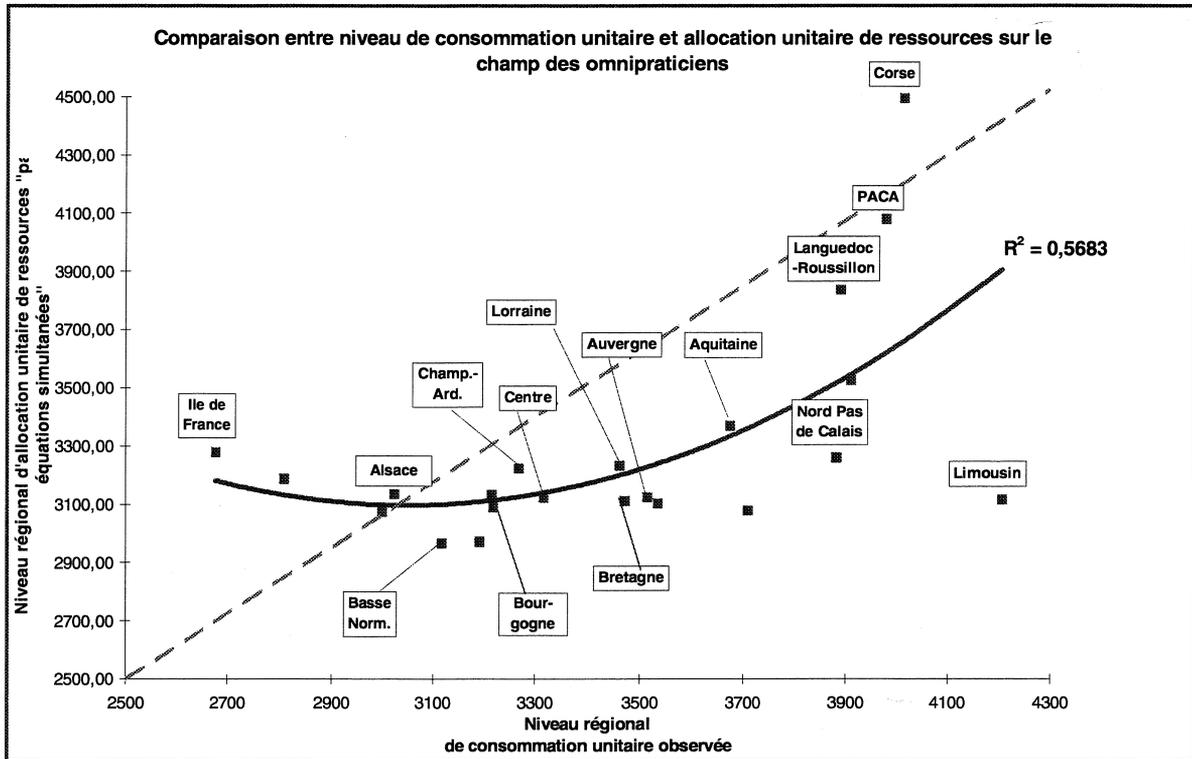
Nous avons retenu la première alternative, car, en réalité, il n'y a, du moins en ce qui concerne les spécialistes, aucune relation linéaire exploitable entre la consommation des ressources et l'indicateur optimal de satisfaction des besoins, alors que cette relation existe dès lors que l'indicateur de satisfaction des besoins est sous-optimal.

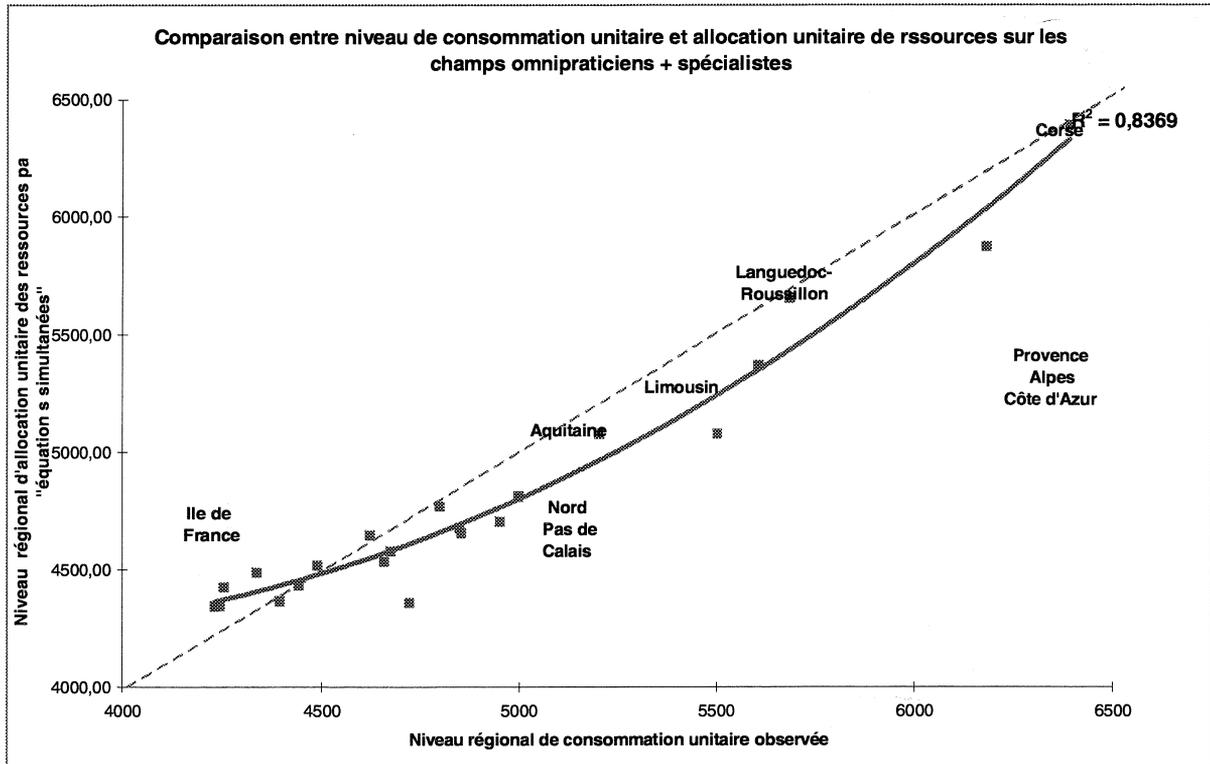
4. Les résultats de la simulation

4.1. Ils sont illustrés par les trois graphiques ci-après, qui appellent les commentaires suivants :

- sur chacun des graphiques la courbe de tendance est celle qui maximise le coefficient de détermination ; il s'agit à chaque fois d'une forme quadratique,
- il a été mis en évidence sur toutes les régions et visualisé plus précisément sur certaines d'entre elles que la procédure d'allocation des ressources choisie comportait un transfert intersectoriel,
- ce résultat est particulièrement démontré par le fait que le coefficient de détermination entre allocation de ressources et consommation de ressources est maximum dès lors qu'il porte sur la somme des deux champs,
- il convient de noter que si la procédure comporte en pratique une redistribution des enveloppes, celle-ci s'effectue significativement de façon privilégiée en direction des spécialistes - environ 15 régions -, contre 3 régions s'agissant des omnipraticiens ; ceci est, en grande partie, la conséquence d'un déséquilibre latent dans l'implantation locale des spécialistes.

Remarque : il eut été préférable que les « t de Student » relatifs aux secondes équations du système (Annexe page 2) ne soient pas inférieurs à 1.73 en raison du nombre limité d'observations. Cet objectif pourrait être atteint en travaillant sur des données départementales (95 observations) plutôt que régionales.





4.2. On peut également tirer un enseignement des coefficients directeurs associés aux facteurs explicatifs retenus dans les équations simultanées.

Ils figurent dans les sorties SAS présentées en Annexe.

Lorsqu'il s'agit des omnipraticiens, le coefficient directeur associé aux spécialistes = 1,006

Lorsqu'il s'agit des spécialistes, le coefficient directeur associé aux omnipraticiens = 0,3926

Si les transferts de consommation entre omnipraticiens et spécialistes s'effectuaient de façon équilibrée, ils devraient être dans le même rapport que celui existant entre les niveaux nationaux de consommation unitaire sur chacun des deux champs

Au plan national ce rapport est très voisin de 2, alors que le rapport entre les coefficients directeurs est sensiblement = 2,6.

Ceci traduit un déséquilibre d'environ 30% au détriment des spécialistes.

Ceci signifie économiquement que les omnipraticiens ont davantage tendance à conserver leurs clients et à les soigner eux-mêmes que les spécialistes, lesquels, de façon plus générale, auront tendance à renvoyer leurs clients vers les omnipraticiens chargés d'en assurer le suivi.

Il faut être conscient de ce phénomène au moment de la mise en place de médecins référents.

ANNEXE

The SAS System

SYSLIN Procedure

First Stage Regression Statistics

Model:

Dependent variable: **YOMN**

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	2	2049588.0022	1024794.0011	11.166	0.0006
Error	19	1743797.4978	91778.81568		
C Total	21	3793385.5000			

Root MSE	302.95019	R-Square	0.5403
Dep Mean	3297.50000	Adj R-SQ	0.4919
C.V.	9.18727		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > T
INTERCEP	1	1651.787145	387.935087	4.258	0.0004
MORBB1	1	839.420348	359.509594	2.335	0.0307
PRECA	1	720.727009	209.891819	3.434	0.0028

Model:

Dependent variable: **YSPE**

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	2	846020.86030	423010.43015	8.670	0.0021
Error	19	927022.95789	48790.68199		
C Total	21	1773043.8182			

Root MSE	220.88613	R-Square	0.4772
Dep Mean	1527.09091	Adj R-SQ	0.4221
C.V.	14.46450		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > T
INTERCEP	1	368.823813	282.850064	1.304	0.2078
MORBB1	1	834.339514	262.124554	3.183	0.0049
PRECA	1	282.968225	153.035692	1.849	0.0801

Two-Stage Least Squares Estimation

Dependent variable: YOMN

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	2	2049588.0022	1024794.0011	5.860	0.0104
Error	19	3322641.0309	174875.84373		
C Total	21	3793385.5000			

Root MSE	418.18159	R-Square	0.3815
Dep Mean	3297.50000	Adj R-SQ	0.3164
C.V.	12.68178		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > T
INTERCEP	1	1280.717324	722.778365	1.772	0.0924
YSPE	1	1.006090	0.594787	1.692	0.1071
PRECA	1	436.035607	368.113983	1.185	0.2508

Durbin-Watson 1.544
 (For Number of Obs.) 22
 1st Order Autocorrelation 0.177

Dependent variable: YSPE

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	2	846020.86030	423010.43015	5.559	0.0126
Error	19	1445769.4174	76093.12723		
C Total	21	1773043.8182			

Root MSE	275.84983	R-Square	0.3692
Dep Mean	1527.09091	Adj R-SQ	0.3027
C.V.	18.06375		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > T
INTERCEP	1	-279.692574	657.072615	-0.426	0.6751
YOMN	1	0.392615	0.265171	1.481	0.1551
MORBB1	1	504.770505	437.527025	1.154	0.2629

Durbin-Watson 1.700
 (For Number of Obs.) 22
 1st Order Autocorrelation 0.109