

VÉRIFICATION EMPIRIQUE DE DEUX
SCHÉMAS D'ANTICIPATION : ADAPTATIF
ET RATIONNEL

C. GOURIEROUX *et* I. PEAUCELLE

N° 8517

- RÉSUMÉ -

Dans cet article on procède à des tests empiriques des anticipations d'évolution d'effectifs faites par les entreprises. On utilise pour ces tests des données de l'"Enquête trimestrielle sur la situation et les perspectives dans l'industrie". On vérifie la validité respectivement du schéma adaptatif et rationnel au niveau agrégé, en distinguant l'activité économique et la taille des entreprises et en tenant compte explicitement de l'aspect qualitatif des réponses individuelles. Dans les deux cas on étudie le problème d'agrégation.

- ABSTRACT -

In this paper we perform empirical tests on the expectations of employment made by the firms. We use qualitative panel data from quarterly tendency surveys in industry. We verify the validity of rational and extrapolative expectation hypotheses on aggregated level, according to the economic and size characteristics of the firms and taking into account explicitly the qualitative aspect of individual responds. We study in both cases the aggregation problem.

1. Introduction

Les modèles existants contiennent un nombre important d'hypothèses a priori et comme le souligne E. Malinvaud, "leur spécification fait intervenir un ensemble de présupposés théoriques, dont la validité reste contestable". Dans les modèles macroéconomiques par exemple la dynamique est souvent introduite par l'intermédiaire des anticipations de certains agrégats. Ces anticipations étant généralement non observables, il est alors nécessaire de préciser la façon dont celles-ci sont obtenues à partir des variables disponibles. Deux schémas d'anticipations sont essentiellement retenus en pratique.

Dans la forme adaptative / extrapolative, la prévision de la variable est exprimée comme une moyenne pondérée des valeurs passées de cette même variable. Une telle formulation est simple à utiliser, mais aussi assez restrictive : les autres variables du système, y compris celles corrélées avec la variable d'intérêt, sont supposées ne pas servir à la prévision.

L'hypothèse d'anticipations rationnelles introduite par Muth (1961) consiste, elle, à supposer que la variable est prévue au mieux en utilisant toute l'information disponible. Il y a donc prise en compte pour prévoir de toutes les variables du système et utilisation optimale des relations existant entre ces variables.

Il n'est cependant pas facile de savoir dans le cadre d'un modèle donné, si l'un de ces schémas est préférable à l'autre. En effet, dans la forme réduite, qui seule est utilisable au niveau de l'estimation, les contraintes impliquées par le schéma d'anticipation sont généralement indistinguables des contraintes provenant du modèle structurel qui lie les variables endogènes aux exogènes et aux anticipations. Les tests faits à partir de modèles sont alors des tests joints sur la forme structurelle et le schéma d'anticipation.

Une autre approche cependant est envisageable. Les diverses modélisations des anticipations peuvent être directement analysées, lorsqu'on dispose de données sur les réalisations de certaines variables et sur leurs prévisions faites par les agents économiques. Cette approche directe qui nécessite l'existence de telles données d'enquête a par exemple été suivie par Turnovsky S., Wachter (1972) ; Carlson J. (1977) ; Mullineaux D. (1978) ; Friedman B. (1980) ; Zimmermann K.F. (1984) ; Ivaldi M. (1985).

Elle présente l'avantage de ne pas s'appuyer sur une formulation du fonctionnement réel de l'économie. Il s'agit donc en un certain sens d'une approche "sans modèle". Elle doit en revanche prendre en compte le caractère microéconomique des données et, par exemple, examiner si la façon d'anticiper dépend ou non des caractéristiques individuelles des agents.

Dans la littérature, cette approche directe a essentiellement été appliquée à l'analyse de l'inflation. Nous avons choisi ici d'étudier la formation des anticipations des entreprises concernant l'évolution de leurs effectifs. L'intérêt économique d'une telle étude est important surtout, si on se réfère comme dans ce travail, à la période 1974-1981. Durant cette période, diverses politiques spécifiques ont en effet été mises en oeuvre pour maintenir l'emploi (dans la perspective d'une reprise de la croissance) ou pour "gérer au mieux" les réductions d'effectifs (dans le cadre du redéploiement industriel). Evaluer l'efficacité éventuelle de telles politiques nécessite de connaître la façon dont les entreprises perçoivent ces modifications de leur environnement, de savoir dans quel sens elles vont réagir et avec quel délai. Une étude détaillée de leurs prévisions à l'embauche, des contraintes sur l'emploi qu'elles pensent devoir subir... peut fournir une information utile sur leur comportement à venir.

Pour l'analyse empirique nous utilisons les données individuelles provenant des enquêtes de conjoncture de l'Institut National de la Statistique et des Etudes Economiques. Elles permettent de constituer des séries de l'évolution des effectifs dans les secteurs d'activités et celles des prévisions faites par les enquêtés à une période antérieure. On obtient ainsi les taux de variation des effectifs réels et anticipés.

La description de l'enquête, de la façon dont sont rédigées les questions et des méthodes usuelles de traitement sont discutées dans le paragraphe 2.

Le schéma adaptatif est présenté dans le paragraphe 3. On étudie dans le cas de données quantitatives l'évolution dans le temps du coefficient d'adaptation au niveau agrégé, puis en distinguant les entreprises par secteur et taille. Une analyse similaire est faite en prenant explicitement en compte l'aspect qualitatif des données.

Les résultats obtenus peuvent être en partie expliqués par des problèmes d'agrégation. Ceux-ci sont étudiés dans la section 4 en supposant successivement la structure constante, puis variant dans le temps. Ceci conduit à expliquer les modifications observées des coefficients d'adaptation par l'intermédiaire de variables exogènes comme l'évolution du chiffre d'affaire.

La même démarche est reprise pour analyser la modélisation par anticipation rationnelle.

Diverses procédures de test de cette hypothèse sont appliquées d'abord aux données agrégées, puis en distinguant les entreprises par secteur et taille. Les résultats obtenus dépendent beaucoup de la définition retenue pour la notion de prévision optimale et de la prise en compte explicite ou non de l'aspect qualitatif des réponses.

Le problème d'agrégation de schémas rationnels est étudié dans le paragraphe 6.

2. Les enquêtes de conjoncture de l'INSEE

La nécessité de mieux connaître les modifications à court terme de l'économie a conduit au début des années soixante à la création d'un système d'enquêtes adaptées à l'analyse conjoncturelle. Ces enquêtes sont en France effectuées sous la responsabilité de l'INSEE, en liaison avec les autres pays de la Communauté Européenne. Elles présentent la particularité de comporter principalement des questions relatives aux tendances : évolution récente de certaines variables et évolutions envisagées pour le futur proche. Elles sont menées auprès des ménages, des entreprises, des établissements financiers, du commerce... et s'intéressent à des sujets comme la consommation, l'activité industrielle, les investissements, la situation de trésorerie, la concurrence étrangère... Parmi ces enquêtes nous avons retenu l'"enquête trimestrielle sur la situation et les perspectives dans l'industrie". Elle comporte en effet diverses questions relatives aux effectifs des entreprises et aux variables pouvant influencer sur leurs évolutions. D'autre part, la périodicité trimestrielle paraît plus adaptée à l'analyse des anticipations sur l'emploi qu'une périodicité mensuelle, pour laquelle existe également un questionnaire plus léger.

Le questionnaire comporte aussi bien des renseignements de caractère micro-économique : chiffre d'affaire, effectifs, évolutions réalisées et prévues des effectifs et des taux de salaire... dans l'entreprise que de caractère macro-économique : évolution des taux de salaire horaire et des prix pour l'ensemble de l'industrie française. La formulation des questions peut induire une réponse quantitative comme : "De combien ont varié les taux de salaire horaire dans votre entreprise : --- %" ; cependant, les questions sont le plus souvent écrites sous forme qualitative : "Effectifs et durée hebdomadaire du travail :

	effectifs	durée du travail
Tendance probable au cours des 3 ou 4 prochains mois :	↗ → ↘	↗ → ↘

Cet aspect qualitatif présente un avantage certain au niveau de la collecte des données : il permet en effet aux enquêtés de répondre sans perte de temps. Il entraîne en revanche diverses difficultés au niveau du traitement statistique. Ainsi, les données qualitatives sont habituellement transformées en données quantitatives par des procédures empiriques, plus ou moins justifiées. Deux d'entre elles sont utilisées dans la suite : la constitution de table de contingence et le calcul de soldes.

Supposons par exemple que nous souhaitions comparer les réalisations et anticipations d'une même variable y . L'échantillon des enquêtés étant suivi dans le temps, ceci est réalisable en utilisant les données sur les réalisations de l'enquête t et les données sur les prévisions de y_t données dans l'enquête précédente $t-1$. On peut alors établir un tableau à double entrée :

prévision réalisation	↗	→	↘	
↗	$p_{+,+,t}$	$p_{+,=,t}$	$p_{+,-,t}$	$p_{+,.,t}$
→	$p_{=,+,t}$	$p_{=,=,t}$	$p_{=,-,t}$	$p_{=.,t}$
↘	$p_{-,+,t}$	$p_{-,=,t}$	$p_{-,-,t}$	$p_{-,.,t}$
	$p_{.,+,t}$	$p_{.,=,t}$	$p_{.,-,t}$	

où l'un des éléments, $p_{++,t}$ par exemple, fournit la proportion d'entreprises ayant anticipé une augmentation qui s'est effectivement réalisée. Une telle table à deux dimensions est peu maniable et on préfère souvent la résumer par l'intermédiaire des soldes. Le solde s_t associé à la prévision s'obtient en calculant la différence entre la proportion de réponses positives (↗) et la proportion de réponses négatives (↘) :

$$s_t = p_{.,+,t} - p_{.,-,t}$$

De manière similaire le solde associé à la réalisation est :

$$s_t^* = p_{+,.,t} - p_{-,.,t}$$

3. Etude descriptive du schéma adaptatif

Des procédures de prévision usuelles d'une variable quantitative y_t consistent à expliquer celle-ci en fonction de son propre passé. Parmi ces techniques auto-projectives le schéma adaptatif est obtenu en choisissant des pondérations décroissant de manière exponentielle. La prévision de y_t faite à $t-1$ est ainsi de la forme :

$$(1) \hat{y}_t^e = (1-\lambda) \sum_{i=0}^{\infty} \lambda^i y_{t-1-i} \quad \text{avec le paramètre d'adaptation contraint}$$

par $0 \leq \lambda \leq 1$. Cette écriture peut être transformée de manière à faire apparaître un mécanisme de correction d'erreurs : la modification des anticipations est proportionnelle à la dernière erreur d'anticipation :

$$(2) \hat{y}_t^e - \hat{y}_{t-1}^e = \lambda (y_{t-1} - \hat{y}_{t-1}^e) \quad 0 \leq \lambda \leq 1$$

Il faut bien noter à ce niveau que l'équivalence entre les deux formulations repose de manière fondamentale sur l'hypothèse d'invariance temporelle du coefficient λ .

Nous nous proposons dans ce chapitre d'examiner, à partir des données d'évolution d'effectifs tirées des enquêtes de conjoncture, si ce schéma est ou non approximativement satisfait. L'étude est essentiellement descriptive. Dans les deux premiers paragraphes nous examinons l'évolution dans le temps

des rapports $\lambda_t = \frac{\hat{y}_t^e - \hat{y}_{t-1}^e}{y_{t-1} - \hat{y}_{t-1}^e}$ calculés d'abord de manière agrégée sur toutes

les entreprises, puis en distinguant celles-ci par secteur et taille. Il apparaît de manière claire que l'hypothèse de constance temporelle du coefficient d'adaptation ne peut être acceptée et que les contraintes $0 \leq \lambda_t \leq 1$ ne sont pas satisfaites. Dans le troisième paragraphe, cette étude est prolongée en considérant la table de contingence à double entrée, qui donne le croisement entre le sens de l'erreur de prévision $y_{t-1} - \hat{y}_{t-1}^e$ et le sens de révision des anticipations $\hat{y}_t^e - \hat{y}_{t-1}^e$. Les conclusions obtenues sont alors beaucoup moins fermes.

3.a. Etude au niveau agrégé

Les soldes correspondant aux évolutions récentes et envisagées des effectifs sont régulièrement publiées dans "Tendances de la conjoncture" depuis 1966. Une difficulté pratique pour utiliser ces données provient de la modification de la périodicité de l'enquête : celle-ci effectuée en Mars, Juin, Novembre jusqu'en 1978 est devenue ensuite trimestrielle et faite en Janvier, Mars, Juin et Octobre. Ceci n'a pas entraîné de modification sensible de la rédaction des questions "tendance probable au cours des 4 prochains mois" remplacée par "tendance probable au cours des 3 ou 4 prochains mois". Cependant, la modification de l'écart séparant deux enquêtes successives devrait, si le schéma adaptatif est correct, conduire à une modification du coefficient λ . De manière à tenir compte de cette difficulté et simultanément d'éliminer des saisonnalités éventuelles, nous avons calculé les coefficients à partir des données relatives à la période Mars - Juin. A ce niveau, il nous faut discuter l'interprétation d'un schéma adaptatif écrit pour les soldes. Les soldes sont uniquement des résumés statistiques et n'ont pas a priori d'interprétation économique directe ; en particulier il n'y a aucune prévision des soldes qui soit effectuée par les agents. Cependant, il est connu que les soldes calculés sont très fortement corrélés aux variables quantitatives sous-jacentes, ici la différence y_t^* entre les nombres de salariés aux dates t et $t-1$. Cette forte corrélation peut être traduite au moyen d'une relation linéaire :

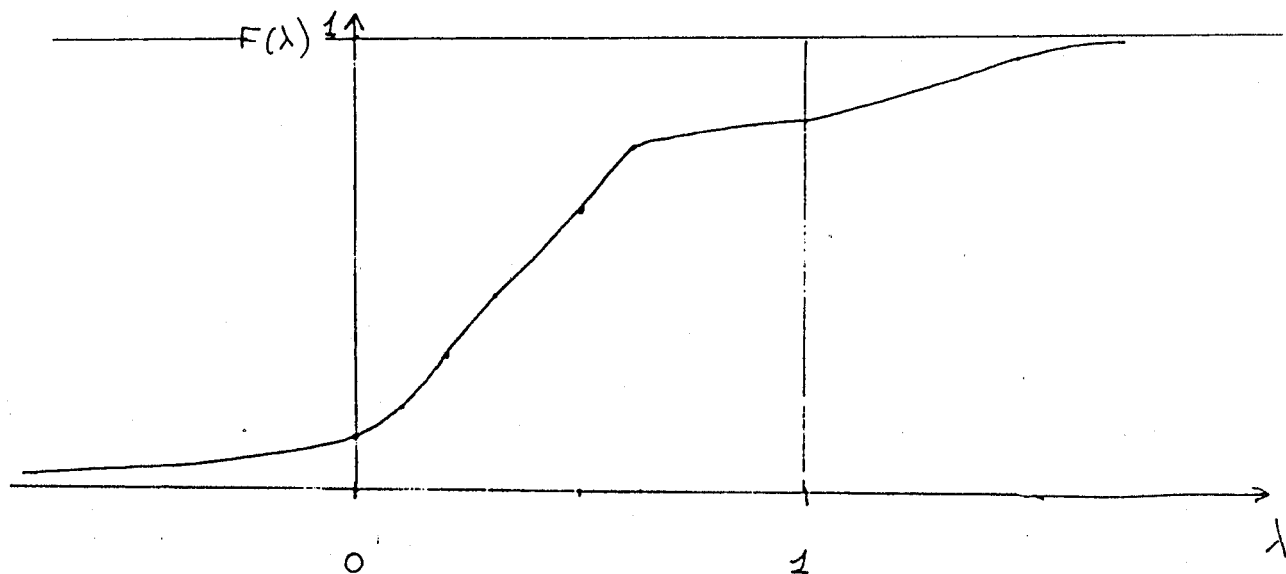
$$\begin{cases} s_t = a + b y_t^* & \text{(relation testée empiriquement)} \\ s_t^e = a + b \hat{y}_t^{*e} & \text{(relation écrite par prolongement de la précédente)} \end{cases}$$

relation qui en pratique apparaît stable dans le temps. Ecrire un schéma adaptatif sur les soldes équivaut alors à l'écrire sur la variable quantitative sous-jacente non observable, puisque :

$$\lambda_t = \frac{s_t^e - s_{t-1}^e}{s_{t-1}^e - s_{t-2}^e} = \frac{\hat{y}_t^{*e} - \hat{y}_{t-1}^{*e}}{y_{t-1}^* - \hat{y}_{t-1}^{*e}}$$

C'est cette interprétation qu'il faut donner aux calculs effectués sur les soldes.

Nous donnons ci-dessous la forme de la distribution empirique des coefficients λ_t calculés à partir des soldes pour l'ajustement des prévisions entre Mars et Juin.



Cette forme de distribution conduit évidemment à rejeter l'hypothèse de constance de λ .

De plus, il est possible à partir de l'ensemble des données d'examiner si la contrainte $0 \leq \lambda \leq 1$ est ou non satisfaite. Nous donnons dans le tableau ci-dessous les proportions de coefficients λ_t respectivement négatifs, compris entre 0 et 1 et supérieurs à 1.

valeur de λ_t	< 0	$0 \leq \lambda_t \leq 1$	> 1
proportions	0.14	0.64	0.22

3.b. Etude par secteur d'activité et taille d'entreprise

L'analyse précédente peut être affinée en prenant en compte certaines caractéristiques des entreprises comme le secteur et la taille. Ceci permettra d'examiner si la non constance du coefficient λ provient simplement de différences de comportements d'anticipation entre secteurs et groupes de taille ou si le problème est plus profond et si la non invariance subsiste à l'intérieur de chacune des sous-populations.

Les entreprises ont été classées selon la nature du produit principal. Quatre secteurs industriels ont été retenus et la répartition moyenne de l'échantillon entre ces secteurs est donnée ci-après.

Secteurs	Nombre moyen d'entreprises
Industries agricoles et alimentaires	200
Industries des biens intermédiaires	750
Industries des biens d'équipement	500
Industries des biens de consommation courante	750

Pour chaque secteur et chaque enquête, on peut déterminer les soldes relatifs aux prévisions et réalisations des évolutions d'effectifs et en déduire les répartitions du coefficient d'adaptation selon les catégories : " $\lambda < 0$ ", " $0 \leq \lambda \leq 1$ ", " $\lambda > 1$ ".

secteur \ valeur de λ	< 0	$0 \leq \leq 1$	> 1	Total
I.A.A.	0.40	0.28	0.32	1
Biens intermédiaires	0.49	0.27	0.24	1
Biens d'équipement	0.41	0.24	0.35	1
Biens de consommation	0.42	0.29	0.29	1

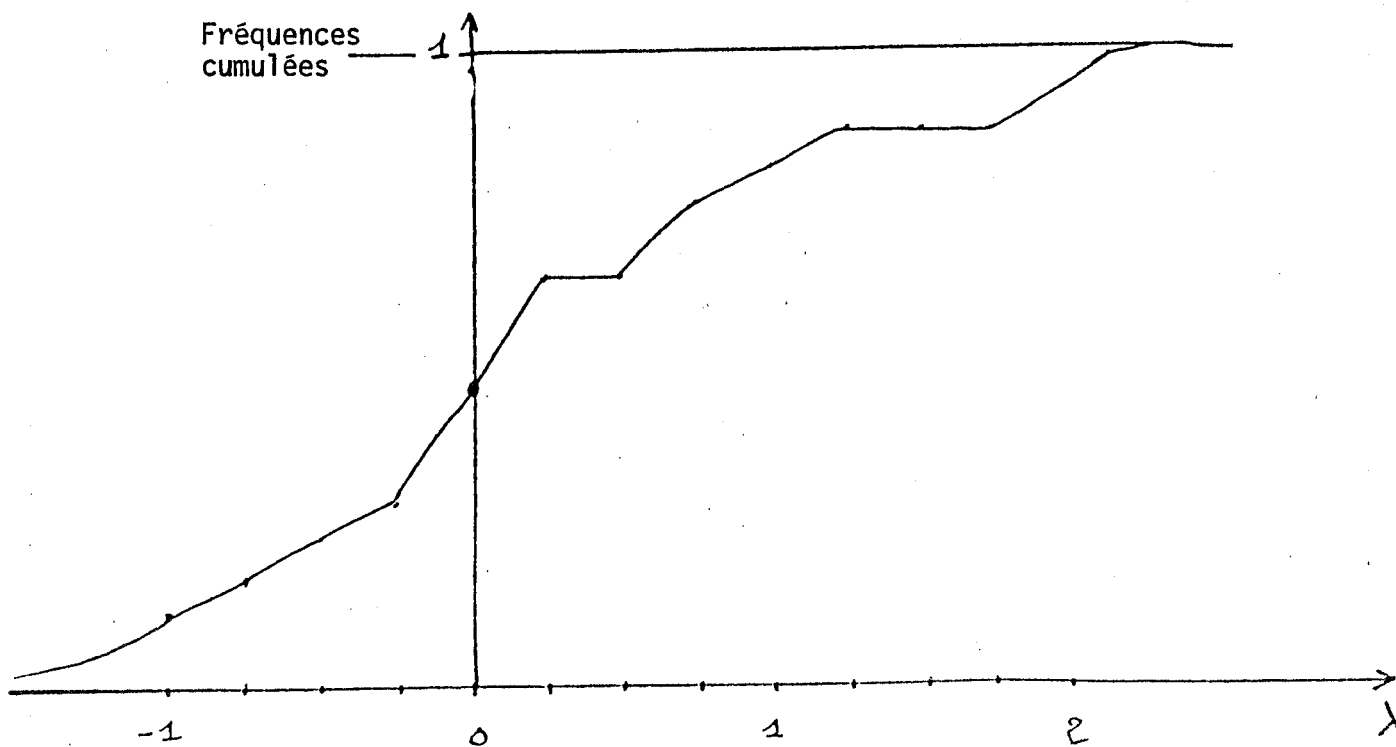
Ces résultats montrent une grande stabilité de la distribution de λ en fonction du secteur d'activité. Surtout ils mettent en évidence une diminution sensible de la proportion d'observations pour lesquelles λ est compris entre 0 et 1 ; dans le cas agrégé (voir 3.a.) la valeur obtenue .64 est deux fois plus élevée (1). Ceci conduirait donc à un rejet plus marqué de l'hypothèse de schéma adaptatif.

(1) L'indépendance de la distribution de λ en fonction du secteur n'implique pas qu'au niveau agrégé la distribution doit être confondue avec celles obtenues au niveau désagrégé. Il est en effet facile de voir en explicitant la procédure d'agrégation (voir § 4) que la distribution agrégée n'est pas une combinaison convexe des distributions désagrégées.

Le même type de conclusion est obtenu, lorsqu'on désagrège les entreprises selon la taille ; cinq classes ont été retenues :

nombre de salariés	valeur de λ			Total
	< 0	$0 \leq \leq 1$	> 1	
< 50	0.48	0.26	0.26	1
$50 \leq < 100$	0.42	0.27	0.31	1
$100 \leq < 500$	0.44	0.21	0.35	1
$500 \leq < 1000$	0.43	0.28	0.29	1
$1000 \leq$	0.40	0.34	0.26	1

Cette répartition presque uniforme du coefficient λ selon les trois modalités implique évidemment une non constance de ce dernier. Celle-ci s'observe même au niveau plus fin de croisements par secteur d'activité et taille. Si nous choisissons par exemple les entreprises de plus de 1000 salariés du secteur des biens intermédiaires, nous pouvons construire la distribution empirique des coefficients λ . Alors que les valeurs possibles de λ construit à partir des soldes sont toutes les valeurs réelles positives ou négatives, les valeurs réellement chargées sont comprises entre -1 et 2.



3.c. Etude directe à partir des tableaux de contingence

La formule (1) définissant le schéma adaptatif étant écrite pour des variables quantitatives, nous avons dans les deux paragraphes précédents suivi la démarche usuelle consistant à rendre quantitatives les données d'enquête. Ceci a alors permis de calculer des valeurs du coefficient d'adaptation λ pour les diverses dates. Cependant une autre approche est envisageable et paraît plus adaptée aux données disponibles. Il est en effet possible à partir des données brutes de déterminer, sinon la valeur de λ , du moins son signe. Ainsi d'observations comme : $\hat{y}_{t-1}^e = "="$, $y_{t-1} = "+"$, $\hat{y}_t^e = "-"$, on peut déduire que le coefficient d'adaptation associé à la variable quantitative sous jacente est négatif. Il y a au total 27 modalités possibles pour le triplet $(\hat{y}_{t-1}^e, y_{t-1}, \hat{y}_t^e)$ et nous donnons dans les tableaux ci-dessous les renseignements sur le signe de λ qui peuvent en être déduits.

Quand $\hat{y}_{t-1}^e = (\nearrow)$

$y_{t-1} \backslash \hat{y}_t^e$	(=) \nearrow	(-) \rightarrow	(-) \searrow
\nearrow (=)	indet	indet	indet
\rightarrow (-)	indet	> 0	> 0
\searrow (-)	indet	> 0	> 0

Quand $\hat{y}_{t-1}^e = (\rightarrow)$

$y_{t-1} \backslash \hat{y}_t^e$	(+) \nearrow	(=) \rightarrow	(-) \searrow
\nearrow (+)	> 0	indet	< 0
\rightarrow (=)	indet	indet	indet
\searrow (-)	< 0	indet	> 0

Quand $\hat{y}_{t-1}^e = (\searrow)$

$y_{t-1} \backslash \hat{y}_t^e$	(+) \nearrow	(+) \rightarrow	(=) \searrow
\nearrow (+)	> 0	> 0	indet
\rightarrow (+)	> 0	> 0	indet
\searrow (=)	indet	indet	indet

Les renseignements sur le signe de λ sont obtenus en comparant la croissance ou décroissance de y_{t-1} par rapport à \hat{y}_{t-1}^e avec celle

de \hat{y}_t^e par rapport à \hat{y}_{t-1}^e . Nous avons pour chaque entrée fait figurer ce renseignement sur le sens de variation (+), (=), (-). Pour la colonne $(y_{t-1}, \hat{y}_{t-1}^e)$, ce n'est autre que la fonction "surprise" introduite par Nerlove M. (1983).

Dès que deux réponses y_{t-1} et \hat{y}_{t-1}^e ou \hat{y}_t^e et \hat{y}_{t-1}^e coïncident l'information n'est pas suffisante pour déterminer le signe du coefficient d'adaptation.

La probabilité :

$$p_- = \Pr [\hat{y}_{t-1}^e = "\rightarrow", y_{t-1} = "\searrow", \hat{y}_t^e = "\nearrow"]$$

$$+ \Pr [\hat{y}_{t-1}^e = "\rightarrow", y_{t-1} = "\nearrow", \hat{y}_t^e = "\searrow"]$$

fournit donc une borne inférieure (borne qui pourrait être atteinte) pour la probabilité que λ soit négatif.

De même p_+ , probabilité obtenue en sommant les probabilités des cases du tableau associées à " > 0 ", fournit une borne inférieure de la probabilité que λ soit positif. Les seuls renseignements concernant le signe du coefficient d'adaptation sont donc résumés par :

$$p_- \leq \Pr [\lambda < 0] \leq 1 - p_+$$

$$p_+ \leq \Pr [\lambda > 0] \leq 1 - p_-$$

Le calcul direct à partir des données individuelles des fréquences empiriques correspondant à p_- et p_+ donne :

$$\hat{p}_- = 1 \% \quad \text{et} \quad \hat{p}_+ = 29 \%$$

Nous voyons donc que :

$29 \% \leq \Pr (\lambda > 0) \leq 99 \%$

Lorsque l'aspect purement qualitatif des données est explicitement pris en compte, il est seulement possible de déterminer un encadrement de $\Pr [\lambda > 0]$; la borne supérieure de celui-ci est très proche de 1 et il apparaît donc difficile de rejeter l'hypothèse $\Pr (\lambda > 0) = 1$. Cette approche adaptée aux données qualitatives nous conduit donc à une conclusion moins ferme que celles obtenues en 3.a. et 3.b. à partir de techniques de type quantitatif ; en particulier l'étude du signe du coefficient d'adaptation ne conduit pas à un rejet de l'hypothèse de schéma adaptatif.

La même démarche peut être suivie pour étudier la position de λ par rapport à 1. Les renseignements sur cette position sont résumés dans les tableaux ci-dessous :

Quand $\hat{y}_{t-1}^e = (\nearrow)$

$y_{t-1} \backslash \hat{y}_t^e$	(=) \nearrow	(-) \rightarrow	(-) \searrow
\nearrow (=)	indet	indet	indet
\rightarrow (-)	indet	indet	> 1
\searrow (-)	indet	< 1	indet

Quand $\hat{y}_{t-1}^e = (\rightarrow)$

$y_{t-1} \backslash \hat{y}_t^e$	(+) \nearrow	(=) \rightarrow	(-) \searrow
\nearrow (+)	indet	indet	indet
\rightarrow (=)	indet	indet	indet
\searrow (-)	indet	indet	indet

Quand $\hat{y}_{t-1}^e = (\searrow)$

$y_{t-1} \backslash \hat{y}_t^e$	(+) \nearrow	(+) \rightarrow	(=) \searrow
\nearrow (+)	indet	< 1	indet
\rightarrow (+)	> 1	indet	indet
\searrow (=)	indet	indet	indet

Ici encore il y a peu de modalités permettant de déduire avec certitude que $\lambda < 1$ ou que $\lambda > 1$.

$$\text{On a : } \Pr [\hat{y}_{t-1}^e = \searrow, y_{t-1} = \nearrow, \hat{y}_t^e = \rightarrow] + \Pr [\hat{y}_{t-1}^e = \nearrow, y_{t-1} = \searrow, \hat{y}_t^e = \rightarrow] = 0.02$$

$$\text{et } \Pr [\hat{y}_{t-1}^e = \nearrow, y_{t-1} = \rightarrow, \hat{y}_t^e = \searrow] + \Pr [\hat{y}_{t-1}^e = \searrow, y_{t-1} = \rightarrow, \hat{y}_t^e = \nearrow] = 0.01$$

$$\text{et donc : } 0.01 \leq \Pr [\lambda > 1] \leq 0.98 .$$

Il est intuitif qu'à partir de données sur un sens d'évolution, c'est-à-dire sur une position par rapport à zéro, il y a peu d'information apportée sur la position d'un paramètre comme λ par rapport à une valeur non nulle. Ceci explique l'importance de la fourchette obtenue. Cet encadrement ne permet pas de rejeter l'hypothèse $\lambda < 1$ et par suite l'hypothèse de schéma adaptatif.

4. Agrégation de schémas adaptatifs

Dans ce chapitre nous nous proposons d'examiner comment se comporte le coefficient d'adaptation, lorsqu'il y a agrégation sur les individus. Nous considérons d'abord le cas d'une structure constante et montrons que le coefficient d'adaptation au niveau agrégé est une moyenne pondérée des coefficients individuels. Ce résultat n'est plus valable dans le cas d'une structure se modifiant avec le temps. En particulier le coefficient agrégé λ peut prendre des valeurs négatives ou supérieures à 1 alors que les coefficients individuels sont tous compris entre 0 et 1. Ceci nous incite alors à expliquer les résultats du chapitre 3 en introduisant une variable explicative qui induit une structure dépendant du temps. La modification du chiffre d'affaire apparaît comme une telle variable pertinente.

4.a. Agrégation à structure constante

Nous nous plaçons dans le cas de données quantitatives et supposons donc que les soldes constituent de bonnes approximations des variables quantitatives sous jacentes. Les entreprises sont partitionnées en $k = 1 \dots K$ sous-populations et on admet que pour chacune de ces sous-populations le schéma adaptatif est considéré comme valable. Le coefficient d'adaptation de la k è sous-population relatif à la date t est :

$$\lambda_{kt} = \frac{\hat{s}_{k,t}^e - \hat{s}_{k,t-1}^e}{s_{k,t-1} - \hat{s}_{k,t-1}^e}$$

Si la répartition de la population dans les K sous-populations est invariante dans le temps et, si nous notons μ_k , $k = 1 \dots K$ le poids de la k è sous-population, les soldes au niveau agrégé sont donnés par :

$$\hat{s}_t^e = \sum_{k=1}^K \mu_k \hat{s}_{k,t}^e, \quad \hat{s}_{t-1}^e = \sum_{k=1}^K \mu_k \hat{s}_{k,t-1}^e, \quad s_{t-1} = \sum_{k=1}^K \mu_k s_{k,t-1}$$

et le coefficient d'adaptation agrégé est :

$$\lambda_t = \frac{\hat{s}_t^e - \hat{s}_{t-1}^e}{s_{t-1} - \hat{s}_{t-1}^e}$$

./.

Il peut s'écrire en fonction des coefficients désagrégés $\lambda_{k,t}$ en remplaçant $\hat{s}_t^e, \hat{s}_{t-1}^e, s_{t-1}$ par leurs expressions. On a :

$$\lambda_t = \frac{\sum_{k=1}^K \mu_k \hat{s}_{k,t}^e - \sum_{k=1}^K \mu_k \hat{s}_{k,t-1}^e}{\sum_{k=1}^K \mu_k s_{k,t-1} - \sum_{k=1}^K \mu_k \hat{s}_{k,t-1}^e}$$

$$\lambda_t = \sum_{k=1}^K \frac{\mu_k (s_{k,t-1} - \hat{s}_{k,t-1}^e)}{\sum_{k=1}^K \mu_k (s_{k,t-1} - \hat{s}_{k,t-1}^e)} \lambda_{kt}$$

Le coefficient λ_t est une moyenne des coefficients λ_{kt} avec des facteurs de somme 1, non forcément positifs. Ceci montre que même si les coefficients λ_{kt} sont compris entre 0 et 1, la même propriété n'est pas toujours satisfaite au niveau agrégé. Remarquons que pour nos données, la fonction "surprise", qui traduit les idées de sous estimation, de sur estimation ou d'absence de biais d'estimation admet la répartition suivante :

fonction surprise	$y_{t-1} > \hat{y}_{t-1}^e$	$y_{t-1} = \hat{y}_{t-1}^e$	$y_{t-1} < \hat{y}_{t-1}^e$
fréquences	0.22	0.54	0.24

Il apparaît donc tout à fait possible d'avoir des pondérations négatives pour certains λ_{kt} .

Dans le chapitre 3, nous avons retenu deux structures secteur et taille qui varient peu dans le temps pour la population des entreprises. La structure varie un peu plus au niveau de l'échantillon du fait qu'il n'est possible de retenir que les répondants à deux enquêtes successives; cependant cette variation est assez faible. D'autre part, pour ce choix de variables, les pondérations sont positives, de sorte qu'effectivement les résultats obtenus en 3 apparaissent peu compatibles avec un schéma adaptatif au niveau désagrégé par secteur et tailles.

./.

4.b. Agrégation à structure variable

Lorsque la structure de la population dépend du temps et est donnée par $\mu_{k,t}$ $k=1 \dots K$ pour l'enquête t , les soldes au niveau agrégé sont habituellement calculés en utilisant les pondérations de la date d'enquête. On a :

$$\hat{s}_t^e = \sum_{k=1}^K \mu_{k,t-1} \hat{s}_{k,t}^e, \quad s_{t-1} = \sum_{k=1}^K \mu_{k,t-1} s_{k,t-1}$$

et
$$\hat{s}_{t-1}^e = \sum_{k=1}^K \mu_{k,t-2} \hat{s}_{k,t-1}^e$$

Le coefficient agrégé λ_t est :

$$\lambda_t = \frac{\hat{s}_t^e - \hat{s}_{t-1}^e}{s_{t-1} - \hat{s}_{t-1}^e} = \frac{\sum_{k=1}^K \mu_{k,t-1} \hat{s}_{k,t}^e - \sum_{k=1}^K \mu_{k,t-2} \hat{s}_{k,t-1}^e}{\sum_{k=1}^K \mu_{k,t-1} s_{k,t-1} - \sum_{k=1}^K \mu_{k,t-2} \hat{s}_{k,t-1}^e}$$

et n'admet plus d'expression simple en fonction des λ_{kt} $k = 1 \dots K$.

Pour comprendre les propriétés de la correspondance qui aux λ_{kt} associe λ_t , nous pouvons considérer le cas de deux sous-populations $K = 2$ et d'une modification extrême de la structure entre les deux dates : $\mu_{1,t-2} = 1$ et $\mu_{1,t-1} = 0$.

Le coefficient λ_t s'écrit :

$$\lambda_t = \frac{\hat{s}_{2,t}^e - \hat{s}_{1,t-1}^e}{s_{2,t-1} - \hat{s}_{1,t-1}^e}$$

Au niveau désagrégé :

$$\lambda_{1t} = \frac{\hat{s}_{1,t}^e - \hat{s}_{1,t-1}^e}{s_{1,t-1} - \hat{s}_{1,t-1}^e} \quad \text{et} \quad \lambda_{2t} = \frac{\hat{s}_{2,t}^e - \hat{s}_{2,t-1}^e}{s_{2,t-1} - \hat{s}_{2,t-1}^e} \quad \text{sont compris entre 0 et 1,}$$

il est clair que lorsque $s_{1,t-1}, s_{2,t-1}, \dots$, varient, le coefficient λ_t est lui susceptible de prendre toute valeur entre $-\infty$ et $+\infty$.

Les résultats obtenus dans le chapitre 3 peuvent donc se révéler compatibles avec des anticipations adaptatives effectuées conditionnellement à une caractéristique évoluant dans le temps (ce qui n'était pratiquement pas le cas du secteur et de la taille).

4.c. Influence de l'évolution du chiffre d'affaire

En plus d'une partition par secteur et taille, nous introduisons maintenant un découpage selon l'évolution du chiffre d'affaire de l'entreprise. Quatre sous-classes sont définies à l'intérieur de chaque sous-population et correspondent aux configurations suivantes :

classe	1	2	3	4
chiffre d'affaire	$CA_t > CA_{t-1}$ $CA_{t-1} > CA_{t-2}$	$CA_t < CA_{t-1}$ $CA_{t-1} < CA_{t-2}$	$CA_t > CA_{t-1}$ $CA_{t-1} < CA_{t-2}$	$CA_t < CA_{t-1}$ $CA_{t-1} > CA_{t-2}$

Les classes 1 et 4 décrivent des tendances respectives à la hausse et à la baisse du chiffre d'affaire. Les deux autres classes contiennent des entreprises pour lesquelles il y a un retournement de la tendance. Remarquons immédiatement que le schéma adaptatif fournit de bonnes prévisions dans un environnement stationnaire, au sens constant dans le temps. Les classes se rapprochant le plus de ce contexte idéal sont les classes 2 et 3 et c'est pour ces classes qu'il faut s'attendre à trouver satisfaite la condition $0 < \lambda < 1$.

En présence de tendances à la baisse (cas 4), le schéma adaptatif conduit à des sur estimations systématiques et de plus en plus fortes. Il faut donc s'attendre à trouver des valeurs du coefficient d'adaptation permettant un certain rattrapage et sortant éventuellement de l'intervalle $[0,1]$. Le même argument est applicable aux tendances à la hausse. Les valeurs obtenues du coefficient λ sont données dans le tableau ci-après et confirment partiellement ceci.

./.

Secteur	Taille	Sous classe			
		1	2	3	4
I.A.A.	< 50	0.13	0.37		4.28
	50-100	0.11			0.74
	100-500	1.54	0.40	0.10	- 1.60
	500-1000	0.30	0.25	0.45	0.93
	> 1000	- 0.45	0.54	0.50	0.99
Biens intermédiaires	< 50	0.10	0.75 *		24.4
	50-100	0.14	0.20		1.70
	100-500	3.90	1.20		0.10
	500-1000	1.50	0.60	0.3	0.80
	> 1000	1.00	0.60	1.20	1.10
Biens d'équipement	< 50	0.70	0.70	0.50	0.04
	50-100	0.30	0.90	0.54	0.70
	100-500	1.10	0.70	1.10	3.60
	500-1000	- 0.20	0.66	0.80	7.70
	> 1000	0.60	0.70		- 0.70
Biens de consommation	< 50	1.30	0.80	0.80	4.60
	50-100	2.30		- 0.25 *	- 1.40
	100-500	- 1.90	0.30	0.70	0.30
	500-1000	0.96	0.75	0.50	2.00
	> 1000	0.40	0.50 *	0.30	0.70

(Les cases vides correspondent aux sous échantillons de taille trop faible).

./.

Les proportions de coefficients appartenant à l'intervalle [0,1] sont respectivement :

Classe	1	2	3	4
Proportion dans [0,1]	0.55	0.94	0.77	0.50

Une analyse plus fine aurait demandé de distinguer, parmi les retournements par exemple, ceux qui sont conjoncturels et ceux qui sont structurels. L'idée de stabilité, par laquelle nous avons pressenti les résultats des classes 2 et 3, repose évidemment sur une interprétation conjoncturelle. Celle-ci semble plus probable avec les données dont on dispose qui présentent certaines saisonnalités. En présence d'une forte tendance les retournements saisonniers seraient estompés. Observer un retournement entre deux trimestres consécutifs reflète donc vraisemblablement une tendance assez faible.

Choisir le découpage des sous-échantillons selon les quatre catégories précédentes revient à expliquer le coefficient d'adaptation en fonction du type d'évolution du chiffre d'affaire. On peut cependant se demander si par exemple la valeur de λ dépend seulement de la présence ou non d'une tendance à la hausse ou si elle est également fonction de l'importance de celle-ci. Nous avons pour chacune des catégories régressé λ sur la valeur de la modification du chiffre d'affaire. Il est apparu que cette dernière n'était que très rarement significative. Les quelques cas, où cette variable est apparue comme réellement explicative sont signalés par une * dans le tableau de la page 17. Les prévisions des évolutions d'effectifs semblent donc principalement fondées sur une connaissance qualitative et non quantitative de l'évolution du chiffre d'affaire.

5. Etude du schéma rationnel

5.a. Le test d'"absence de biais"

Une prévision \hat{y}_t^e d'une variable quantitative y_t est dite rationnelle, si elle coïncide avec la meilleure prévision possible fonction de l'information I_{t-1} disponible. Généralement cette anticipation optimale est choisie égale à l'espérance conditionnelle de y_t sachant I_{t-1} et la condition de rationalité s'écrit :

$$\hat{y}_t^e = E(y_t / I_{t-1})$$

Il résulte des propriétés usuelles de l'espérance conditionnelle que :

i) la prévision est sans biais :

$$E \hat{y}_t^e = E E (y_t / I_{t-1}) = E y_t ;$$

ii) la prévision est orthogonale à l'erreur de prévision :

$$E \hat{y}_t^e (y_t - \hat{y}_t^e) = 0$$

Il est facile en pratique de tester ces deux conséquences de l'hypothèse d'anticipations rationnelles. Considérons en effet le modèle de régression :

$y_t = \alpha + \beta \hat{y}_t^e + u_t$, où l'erreur u_t est de moyenne nulle et est orthogonale à \hat{y}_t^e . Les coefficients de régression théorique α et β sont définis par :

$$\beta = \frac{\text{Cov}(y_t, \hat{y}_t^e)}{V(\hat{y}_t^e)} \text{ et } \alpha = E y_t - \beta E(\hat{y}_t^e)$$

Il est facile de voir que les deux conditions :

$$E y_t = E \hat{y}_t^e \text{ et } E \hat{y}_t^e (y_t - \hat{y}_t^e) = 0$$

sont équivalentes à : $\alpha = 0, \beta = 1$

Le test de Fisher de l'hypothèse nulle $H_0 = \{\alpha = 0, \beta = 1\}$ effectué à partir du modèle linéaire : $y_t = \alpha + \beta \hat{y}_t^e + u_t$ est dans la littérature appelé "test d'absence de biais".

Insistons sur le fait que ce test concerne une conséquence de l'hypothèse d'anticipation rationnelle et non cette hypothèse elle-même.

Cependant son usage étant général, il a été appliqué dans la suite sur les observations des soldes.

5.b. Test à partir des soldes

Comment interpréter les conditions d'absence de biais et d'orthogonalité écrites sur les soldes? Nous pouvons pour répondre à cette question suivre le même raisonnement que celui tenu en 3-a pour le schéma adaptatif. Admettant une relation linéaire entre soldes et variables quantitatives sous-jacentes :

$$\begin{cases} s_t = a + b y_t^* \\ \hat{s}_t^e = a + b \hat{y}_t^{*e} \end{cases},$$

nous voyons que la condition d'absence de biais sur les soldes est :

$$\begin{cases} E(s_t - \hat{s}_t^e) = 0 \\ E \hat{s}_t^e (s_t - \hat{s}_t^e) = 0 \end{cases} \Leftrightarrow \begin{cases} b E(y_t^* - \hat{y}_t^{*e}) = 0 \\ ab E(y_t^* - \hat{y}_t^{*e}) + b^2 E \hat{y}_t^{*e} (y_t^* - \hat{y}_t^{*e}) = 0 \end{cases}$$

$$\Leftrightarrow \begin{cases} E(y_t^* - \hat{y}_t^{*e}) = 0 \\ E \hat{y}_t^{*e} (y_t^* - \hat{y}_t^{*e}) = 0 \end{cases}$$

Elle apparaît donc équivalente aux conditions de sans biais et d'orthogonalité écrites pour la variable quantitative sous-jacente.

Ce test dit d'"absence de biais", a été effectué pour l'ensemble des entreprises. La regression estimée de $y_t = s_t$ sur $\hat{y}_t^e = \hat{s}_t^e$ est :

$$y_t = -0.08 + 0.39 \hat{y}_t^e$$

(0.01) (0.09)

Elle conduit au rejet de l'hypothèse nulle.

La même approche peut être suivie en partitionnant l'ensemble des entreprises par secteur et par taille.

Nous donnons ci-dessous les probabilités de dépasser la valeur de la statistique de Fisher de l'hypothèse $\alpha = 0, \beta = 1$. A ce niveau plus désagrégé, il n'y a plus systématiquement rejet de l'hypothèse.

Celle-ci est acceptée avec un risque d'erreur de 5 %, lorsque la probabilité figurant dans le tableau dépasse cette valeur 5 %. Ceci se produit par exemple pour les grandes entreprises (plus de 500 employés) des secteurs "biens d'équipements", "biens intermédiaires", "biens de consommation"

Secteur	Taille	Probabilité de dépasser la valeur de la statistique de Fisher	Refus (R) ou acceptation (A) de l'hypothèse
I.A.A.	< 50	0.01%	R
	50-100	0.3 %	R
	100-500	0.3 %	R
	500-1000	1 %	R
	> 1000	0.01%	R
Biens intermédiaires	< 50	0.01%	R
	50-100	3 %	R
	100-500	19 %	A
	500-1000	31 %	A
	> 1000	15 %	A
Biens d'équipement	< 50	29 %	A
	50-100	0.05%	R
	100-500	2 %	R
	500-1000	12 %	A
	> 1000	8 %	A
Biens de consommation	< 50	0.01%	R
	50-100	15 %	A
	100-500	46 %	A
	500-1000	9 %	A
	> 1000	14 %	A


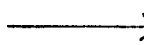



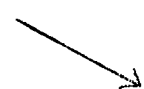
L'hypothèse dit d'"absence de biais", qui a été considérée est caractérisée par les deux conditions: sans biais $E\hat{y}_t^e = Ey_t$ et orthogonalité aux erreurs de prévision: $E\hat{y}_t^e (y_t - \hat{y}_t^e) = 0$. Lorsque cette hypothèse est refusée, il est important d'en déterminer la cause et notamment de savoir si ceci est dû au rejet de la condition de sans biais et/ou au rejet de la condition d'orthogonalité. Signalons d'autre part que si la notion d'anticipation optimale est définie non pas comme l'espérance conditionnelle de y_t sachant I_{t-1} , mais comme la régression linéaire de y_t sachant I_{t-1} , il a été montré (voir Gouriéroux-Pradel (1985)) que l'hypothèse H_0^* = { il existe un ensemble d'information tel que \hat{y}_t^e puisse être considéré comme la regression linéaire de y_t sur cette information } est équivalente à la seule condition d'orthogonalité : $E\hat{y}_t^e (y_t - \hat{y}_t^e) = 0$. Fonder un test sur cette seule condition permet donc aussi de vérifier une notion de rationalité, en un sens il est vrai un peu plus large. Ce test de H_0^* effectué au niveau désagrégé montre que les refus de l'hypothèse H_0 proviennent essentiellement d'un rejet de la condition de sans biais et non de celui de la condition d'orthogonalité.

Test de la condition d'orthogonalité : R (refus), A (acceptation)

Secteur \ Taille	Taille				
	< 50	50-100	100-500	500-1000	> 1000
I.A.A.	A	R	A	A	R
Biens intermédiaires	A	A	A	A	A
Biens d'équipement	A	A	A	A	A
Biens de consommation	A	A	A	A	A

5.c.Prise en compte de l'aspect qualitatif

Le test dit "d'absence de biais", est construit pour des données quantitatives et est donc peut être mal adapté aux données, dont on dispose, qui sont purement qualitatives. Si nous prenons explicitement en compte cet aspect, les conclusions concernant l'hypothèse d'anticipation rationnelle peuvent être sensiblement modifiées. Une procédure de test adaptée au cas qualitatif a été proposée par Gourieroux-Pradel (1985) . Elle est fondée sur l'examen des tables de contingence croisées en y_t, \hat{y}_t^e . L'hypothèse d'anticipation rationnelle s'exprime en effet de manière simple à partir de la loi du couple y_t, \hat{y}_t^e :

$y_t \backslash \hat{y}_t^e$			
	$p_{+,+,t}$	$p_{+,=,t}$	$p_{+,-,t}$
	$p_{=,+,t}$	$p_{=,=,t}$	$p_{=-,t}$
	$p_{-,+,t}$	$p_{-,=,t}$	$p_{-,-,t}$

Elle s'écrit :

$$H_0 = \left\{ \begin{array}{l} p_{++,t} \geq \text{Max} (p_{=,+,t} , p_{-,+,t}), \\ p_{=,=,t} \geq \text{Max} (p_{+,=,t} , p_{-,=,t}), \\ p_{-,-,t} \geq \text{Max} (p_{+,-,t} , p_{=-,t}) \end{array} \right\}$$

Dans une optique descriptive, il suffit donc de construire les tables de contingence empiriques et de regarder si les éléments de la diagonale sont ou non souvent supérieurs aux éléments de la colonne correspondante.

./.

Ces tableaux sont donnés pour l'ensemble des entreprises, pour le classement en secteur et pour le classement par taille; les éléments sont des pourcentages.

Ensemble des entreprises

$y_t \backslash \hat{y}_t^e$			
	1.0	7.6	1.9
	5.1	47.9	12.0
	1.8	16.9	5.7

Classement par secteurs :

I.A.A.

$y_t \backslash \hat{y}_t^e$			
	0.3	4.9	0.7
	4.1	66.6	9.3
	0.9	11.6	1.7

Biens intermédiaires

$y_t \backslash \hat{y}_t^e$			
	0.8	6.9	1.7
	4.8	47.8	12.6
	1.6	17.3	6.5

Biens d'équipement

$y_t \backslash \hat{y}_t^e$			
	2.0	9.8	3.4
	6.6	36.7	12.3
	3.0	18.9	7.3

Biens de consommation

$y_t \backslash \hat{y}_t^e$			
	0.7	7.4	1.4
	4.7	51.6	11.7
	1.4	16.3	4.6

Classement par taille

< 50

$y_t \backslash \hat{y}_t^e$	\nearrow	\rightarrow	\searrow
\nearrow	0.3	4.9	0.5
\rightarrow	3.9	67.7	8.6
\searrow	0.7	11.5	1.9

50-100

$y_t \backslash \hat{y}_t^e$	\nearrow	\rightarrow	\searrow
\nearrow	0.7	7.3	1.1
\rightarrow	5.0	57.9	10.1
\searrow	1.3	14.2	2.4

100-500

$y_t \backslash \hat{y}_t^e$	\nearrow	\rightarrow	\searrow
\nearrow	1.1	8.6	1.9
\rightarrow	5.6	48.4	11.5
\searrow	1.7	16.7	4.6

500-1000

$y_t \backslash \hat{y}_t^e$	\nearrow	\rightarrow	\searrow
\nearrow	1.3	9.0	2.6
\rightarrow	5.5	36.5	14.4
\searrow	2.7	19.8	8.1

> 1000

$y_t \backslash \hat{y}_t^e$	\nearrow	\rightarrow	\searrow
\nearrow	1.3	6.2	3.5
\rightarrow	4.7	30.0	16.1
\searrow	2.9	22.2	13.0

Nous avons entouré les éléments des diagonales correspondant aux colonnes en faveur de l'hypothèse d'anticipations rationnelles. Celles-ci sont en fait obtenues toujours pour la même modalité "→" de la prévision.

Il est ici vraisemblable que les résultats sont en partie masqués par un problème de collecte des données. Les questionnaires ne comporte en effet pour chaque variable que les trois modalités →, →, → et par leur constructi ne permettent pas de fournir une réponse du type "je ne sais pas". Les entrep ses ne connaissant pas de manière très précise l'évolution passée de y ou son évolution anticipée ont alors tendance à répondre : "stabilité". Cette effet de collecte entraîne une surestimation des éléments des deuxièmes lignes et deuxièmes colonnes et de ce fait biaise le test de rationalité vers le rejet pour les premières et troisièmes colonnes.

6. Agrégation de schémas rationnels

6.a. Effet de la structure

Nous nous plaçons dans le cas de données quantitatives et, comme nous l'avions fait pour le schéma adaptatif, nous allons examiner le problème d'agrégation à structure constante et à structure se modifiant dans le temps.

i) Structure constante

Considérons K sous-populations, d'importances respectives μ_k $k = 1 \dots K$. Supposons qu'au niveau désagrégé l'hypothèse d'anticipation rationnelle soit satisfaite :

$$\forall k = 1 \dots K ; \quad \hat{y}_{k,t}^e = E (y_{k,t} / I_{t-1})$$

On en déduit immédiatement qu'elle l'est au niveau agrégé, puisque :

$$\sum_{k=1}^K \mu_k \hat{y}_{k,t}^e = E \left(\sum_{k=1}^K \mu_k y_{k,t} / I_{t-1} \right)$$

$$\hat{y}_t^e = E (y_t / I_{t-1})$$

ii) Structure se modifiant dans le temps

Le résultat précédent ne se généralise évidemment pas puisque :

$$\hat{y}_t^e = \sum_{k=1}^K \mu_{k,t-1} \hat{y}_{k,t}^e$$

et
$$y_t = \sum_{k=1}^K \mu_{k,t} y_{k,t}$$

Un rejet de l'hypothèse d'anticipation rationnelle au niveau agrégé peut donc se révéler compatible avec cette hypothèse écrite à un niveau plus fin.

iii) Prise en compte du chiffre d'affaires

Ceci nous incite à partitionner les entreprises en fonction d'une caractéristique variant dans le temps. Par analogie avec le cas du schéma adaptatif et pour faciliter les comparaisons, nous avons retenu le type d'évolution du chiffre d'affaire. Les résultats des tests d'absence de biais: "R" : rejet , "A" : acceptation, sont résumés dans le tableau ci-dessous :

Test de Fisher de l'hypothèse "d'absence de biais"

Secteur	Taille	Sous classe C.A.			
		1	2	3	4
I.A.A.	< 50	R	R		R
	50-100	R			R
	100-500	R	R	A	R
	500-1000	A	A	R	R
	> 1000	R	A	A	R
Biens intermédiaires	< 50	R	A		R
	50-100	R	A		R
	100-500	R	R	R	A
	500-1000	R	R		A
	> 1000	A	R	A	R
Biens d'équipement	< 50	A	R	A	A
	50-100	A	R	R	R
	100-500	A	R	R	R
	500-1000	A	A	A	R
	> 1000	R	R	A	R
Biens de consommation	< 50	R	A	R	R
	50-100	R	A	R	A
	100-500	R	A	A	A
	500-1000	R	A	R	R
	> 1000	A	R	R	R

(Les cases vides correspondent aux sous échantillons de taille trop faible)
 Un tableau analogue peut être construit pour la sous-hypothèse de non corrélation entre prévision et erreur de prévision.

Test de l'hypothèse d'orthogonalité entre \hat{y}_t^e et $y_t - \hat{y}_t^e$

Secteur	Taille	Sous classe C.A.			
		1	2	3	4
I.A.A.	< 50	A	A		R
	50-100	R			R
	100-500	A	A	A	A
	500-1000	A	A	A	A
	> 1000	R	R	A	R
Biens intermédiaires	< 50	A	A		R
	50-100	R	A		A
	100-500	A	R	A	R
	500-1000	A	R		A
	> 1000	R	A	A	A
Biens d'équipement	< 50	A	A	A	A
	50-100	A	R	R	A
	100-500	A	R	R	A
	500-1000	A	A	A	A
	> 1000	A	R	A	A
Biens de consommation	< 50	R	A	R	R
	50-100	A	A	A	A
	100-500	A	A	A	A
	500-1000	R	A	R	A
	> 1000	A	R	R	A

Comme dans le cas des seules partitions par secteur ou taille, on voit que cette sous-hypothèse est beaucoup plus souvent acceptée (68 % d'acceptation) que l'hypothèse jointe "d'absence de biais" (48 % d'acceptation). On remarque également que les deux tests se révèlent compatibles au sens où, excepté dans deux cas, un rejet de l'hypothèse d'orthogonalité entraîne un rejet de l'hypothèse jointe et une acceptation de l'hypothèse jointe entraîne une acceptation de la sous-hypothèse.

6.b. Effet d'informations différentes

Dans la procédure d'agrégation à structure constante décrite en 6.a, nous avons implicitement supposé que les agents des diverses sous-populations fondaient leurs anticipations sur la même information I_t . Cette hypothèse peut être levée. Dans ce cas l'hypothèse d'anticipation rationnelle au niveau désagrégé s'écrit :

pour tous les $k = 1 \dots K$: $\hat{y}_{k,t}^e = E(y_{k,t} / I_{k,t-1})$

Même à structure constante, elle n'implique plus l'hypothèse d'anticipation rationnelle pour la variable agrégée \hat{y}_t^e . Nous pouvons cependant examiner les conséquences d'informations différentes sur les valeurs α et β des coefficients de la régression :

$$y_t = \alpha + \beta \hat{y}_t^e + u_t$$

De manière à simplifier la présentation, nous supposons qu'il est possible de ranger les informations en ordre croissant :

$$I_{1,t-1} \subset I_{2,t-1} \dots \subset I_{k,t-1}$$

On a alors les égalités suivantes :

$$E \hat{y}_{k,t}^e = E y_{k,t}$$

(condition de sans biais)

$$E \hat{y}_{k,t}^e (y_{k,t} - \hat{y}_{k,t}^e) = 0$$

(orthogonalité entre prévision et erreur de prévision)

$$E [\hat{y}_{k,t}^e (y_{\ell,t} - \hat{y}_{\ell,t}^e)] = 0$$

pour $\ell \geq k$ (croissance des informations)

De ces égalités, on déduit que les prévisions au niveau agrégé sont sans biais, puisque :

$$E \hat{y}_t^e = E \left(\sum_{k=1}^K \mu_k \hat{y}_{k,t}^e \right) = \sum_{k=1}^K \mu_k E(y_{k,t}) = E y_t$$

D'autre part, on a :

$$\begin{aligned} E \hat{y}_t^e (y_t - \hat{y}_t^e) &= E \left[\sum_{k=1}^K \mu_k \hat{y}_{k,t}^e \sum_{\ell=1}^K \mu_{\ell} (y_{\ell,t} - \hat{y}_{\ell,t}^e) \right] \\ &= \sum_{k=1}^K \sum_{\ell=1}^K \mu_k \mu_{\ell} E [\hat{y}_{k,t}^e (y_{\ell,t} - \hat{y}_{\ell,t}^e)] \\ &= \sum_{k>\ell} \sum \mu_k \mu_{\ell} E [\hat{y}_{k,t}^e (y_{\ell,t} - \hat{y}_{\ell,t}^e)] \end{aligned}$$

./.

Cette quantité peut a priori être positive ou négative. Comme α et β sont tels que :

$$\begin{cases} E y_t = \alpha + \beta E \hat{y}_t^e \\ \beta = \frac{\text{cov}(y_t, \hat{y}_t^e)}{V(\hat{y}_t^e)} \end{cases}$$

, on voit en utilisant la condition

de sans biais que :

$$\begin{cases} \alpha = -(\beta-1) E \hat{y}_t^e \\ \beta-1 = \frac{E[\hat{y}_t^e (y_t - \hat{y}_t^e)]}{V(\hat{y}_t^e)} \end{cases}$$

Si à la date t , on prévoit en moyenne une baisse des effectifs ($E \hat{y}_t^e > 0$), le coefficients α et $\beta-1$ devraient être de signes opposés. Cette condition rempli les contraintes $H_0 : (\alpha=0; \beta-1=0)$ obtenus dans le cas d'informations identiques. Il est même possible dans certains cas d'avoir une idée du signe des coefficients. Supposons que les évolutions $y_{k,t}$ soient peu dispersées entre les sous populations. On peut alors considérer que ces valeurs sont confondues avec le résultat agrégé y_t et les écarts entre les prévisions proviennent non pas de réalités différentes, mais seulement d'informations différentes. On a alors si $k > \ell$:

$$\begin{aligned} & E [\hat{y}_{k,t}^e (y_{\ell,t} - \hat{y}_{\ell,t}^e)] \\ &= E \{ E(y_t / I_{k,t}) [y_t - E(y_t / I_{\ell,t})] \} \\ &= E [E(y_t / I_{k,t}) - E(y_t / I_{\ell,t})] [y_t - E(y_t / I_{\ell,t})] \\ &= E E\{ [E(y_t / I_{k,t}) - E(y_t / I_{\ell,t})] [y_t - E(y_t / I_{\ell,t})] / I_{k,t} \} \\ &= E [E(y_t / I_{k,t}) - E(y_t / I_{\ell,t})]^2 \geq 0 \end{aligned}$$

Il faut donc s'attendre à trouver, lorsqu'il y a une faible dispersion des valeurs $y_{k,t}$, une valeur positive de $E \hat{y}_t^e (y_t - \hat{y}_t^e)$, c'est-à-dire des coefficients α et β , qui en période de baisse des effectifs, satisferaient : $\alpha < 0$ et $\beta > 1$. Ces contraintes peuvent être examinées à partir des coefficients de régression estimés dans le modèle $y_t = \alpha + \beta \hat{y}_t^e + u_t$

Ces quelques remarques peuvent être illustrées à partir des données sur les effectifs.

Plaçons nous d'abord au niveau désagrégé par secteur et taille. On peut raisonnablement penser que à secteur donné plus la taille de l'entreprise est importante plus elle détient d'information. Ceci nous conduit à construire pour chaque secteur un tableau à double entrée dont le terme général d'indice k, ℓ est la moyenne empirique des produits croisés entre prévisions

faites par les entreprises de taille k et erreurs de prévision faites par les entreprises de taille ℓ :

$$M_{k\ell} = \frac{1}{T} \sum_{t=1}^T \hat{y}_{k,t}^e (y_{\ell,t} - \hat{y}_{\ell,t}^e)$$

Cette quantité constitue une bonne approximation des moments théoriques $E \hat{y}_{k,t}^e (y_{\ell,t} - \hat{y}_{\ell,t}^e)$. Elle devrait donc apparaître non significative pour $\ell \geq k$, c'est-à-dire lorsque l'élément se trouve dans la partie triangulaire supérieure du tableau. Dans les tableaux ci-dessous, nous avons fait figurer les éléments $M_{k\ell}$ significativement différents de zéro et nous avons remplacé les autres par zéro :

Secteur I.A.A.

$k \backslash \ell$	< 50	50-100	100-500	500-1000	> 1000
< 50	0	0	0	0	0
50-100	0.002	- 0.004	0	0	0
100-500	0	0	- 0.004	0	0
500-1000	0.002	0	0.004	0	0
> 1000	0	0	0	0	- 0.009

Secteur biens intermédiaires

$k \backslash \ell$	< 50	50-100	100-500	500-1000	> 1000
< 50	0	0	0	0.004	0.007
50-100	0	- 0.003	0	0	0
100-500	0	0	0	0.005	0.007
500-1000	0	0	0	0	0
> 1000	0	0	0	0.01	0

Secteur biens d'équipement

k \ ℓ	< 50	50-100	100-500	500-1000	> 1000
< 50	0	0	0	0	0
50-100	0	- 0.004	0	0	0
100-500	0	0	0	0	0
500-1000	0	0	0	0	0
< 1000	0	0	0	0	0

Secteur biens de consommation

k \ ℓ	< 50	50-100	100-500	500-1000	> 1000
< 50	0	0	0	0	0
50-100	0.001	0	0.003	0	0.005
100-500	0	0	0	0	0
500-1000	0	0.003	0.006	0	0.008
> 1000	0	0.002	0.004	0.004	0.006

Le fait que la plupart des éléments soient non significatifs dans le secteur des biens d'équipement montre que l'hypothèse d'informations égales, indépendantes de la taille, serait acceptée. L'hypothèse d'information croissante avec la taille apparaît applicable aux secteurs I.A.A. et biens de consommation. Cette hypothèse est d'ailleurs renforcée par le fait que les éléments non significatifs se trouvant dans les parties triangulaires inférieures sont positifs, ce qui était un résultat attendu.

De manière plus curieuse le tableau associé au secteur biens intermédiaires est lui compatible avec des ensembles d'information décroissant avec la taille.

Finalement nous avons effectué les régressions de y_t sur \hat{y}_t^e pour les périodes correspondant à une baisse prévue des effectifs. Nous donnons ci-dessous les valeurs des estimations de α et β , qui d'après les raisonnements intuitifs faits précédemment, devraient être tels que : $\alpha < 0$, $\beta > 1$ ou au moins tels que α et $\beta-1$ soient de signes opposés. Ces valeurs sont fournies pour les classes, où le nombre d'observations est suffisamment grand.

Secteur	taille	α	β
I.A.A.	100 - 500	- 0.05	1.10
Biens intermédiaires	50 - 100	0.002	- 0.19
Biens d'équipement	< 50	- 0.23	4.20
	50 - 100	- 0.12	0.13
	100 - 500	0.07	- 0.42
	500 - 1000	0.04	- 0.43

L'examen du tableau montre clairement que si la condition α , $\beta-1$ de signes opposés apparaît satisfaite, il n'en est pas de même pour la position de α par rapport à 0 et de β par rapport à 1.

7. Conclusion

Face aux controverses concernant la pertinence de l'hypothèse d'anticipations rationnelles, il nous a paru important d'étudier celle-ci à partir de données directes sur les anticipations d'évolution d'effectifs effectuées par les entreprises et de comparer cette modélisation au schéma plus classique que constitue le modèle adaptatif. Nous avons essentiellement mené une analyse descriptive des données. Il est apparu qu'aucune des deux modélisations n'était acceptée ou refusée de manière systématique. De plus les procédures de tests apparaissent peu robustes. Elles sont en particulier très sensibles à divers traitements statistiques usuels comme l'agrégation des données ou la conversion sous forme quantitatives des réponses qualitatives des agents. Il paraît important d'effectuer de telles études directes à un niveau assez fin, de prendre en compte le caractère qualitatif des données, la possibilité d'information différentes selon les agents, la non constance dans le temps des structures des entreprises et de leur schéma d'anticipation.

- BIBLIOGRAPHIE -

CARLSON J. (1977)

A study of price forecasts, *Annals of Economic and Social Measurement*,
6, 27 - 56

FRIEDMAN B. (1980)

Survey evidence on the rationality of interest rate expectations,
Journal of Monetary Economics, 6, 453 - 465

GOURIEROUX C., PRADEL J. (1985)

Direct test of the rational expectation hypothesis (with special attention
to qualitative variables), à paraître dans *European Economic Review*

MULLINEAUX D. (1978)

On testing for rationality : another look at the Livingstone price
expectations data, *Journal of Political Economy* 86, 329 - 336

IVALDI M. (1985)

On testing rationality of demand expectations from qualitative survey
responses, *Séminaire d'Econométrie (Dir. E. MALINVAUD)*

MALINVAUD E. (1982)

Théorie macro-économique - T.2. - Evolutions conjoncturelles
DUNOD

MUTH J.F. (1961)

Rational expectations and the theory of Price movements, *Econometrica*,
Vol. 29, n° 6, 315 - 335

NERLOVE M. (1983)

Expectations, plans, and realizations in theory and practice,
Econometrica, Vol. 51, n° 5, 1251 - 1279

TURNOVSKY S., WACHTER M. (1972)

A test of the expectations hypothesis using directly observed wage and
price expectations, *Review of Economics and Statistics*, 54, 47 - 54.

ZIMMERMANN K.F. (1984)

On rationality of business Expectations : a micro analysis of qualita-
tive responses, *Séminaire d'Econométrie (Dir. E. MALINVAUD)*.