

CHAPITRE 1

UNE RELATION D'AGENCE AU SEIN DU SERVICE PUBLIC HOSPITALIER

INTRODUCTION

L'évolution des modes de financement des hôpitaux publics et des hôpitaux Privés participant au Service Public Hospitalier (P.S.P.H.), du prix de journée aux ordonnances de 1996, a progressivement contractualisé la vie de ces établissements de santé. Les ordonnances hospitalières de 1996 établissent un cadre relationnel entre l'hôpital, les autorités sanitaires régionales comme l'Agence Régionale de l'Hospitalisation (A.R.H.), la Caisse Régionale d'Assurance Maladie (C.R.A.M.), le Ministère de la Santé. La rémunération de ces hôpitaux, basée sur la Dotation Globale de Fonctionnement (D.G.F.) modulée par l'activité théorique mesurée par le Programme de Médicalisation du Système d'Information (P.M.S.I.), engendre de multiples interactions stratégiques entre les acteurs du système de santé français. Il devient, dès lors, intéressant d'analyser ces interrelations sous l'angle de la théorie normative de l'agence. Après une présentation des réformes du mode de financement des hôpitaux (section 1), nous passons en revue, à partir de la littérature économique de la théorie normative de l'agence, les modèles d'agence (section 2). Enfin, nous proposons un modèle hiérarchique d'agence applicable aux ordonnances de 1996 (section 3).

SECTION 1 : LES RÉFORMES DU MODE DE FINANCEMENT DU SERVICE PUBLIC HOSPITALIER

Critiqué pour son caractère inflationniste dans un contexte de fort ralentissement de la croissance économique, le prix de journée est officiellement remplacé par la D.G.F. à partir de 1983. Ce nouveau mode de financement des hôpitaux publics et P.S.P.H., sans être remis en cause, est réaménagé au cours des années 1990, par l'application du P.M.S.I. et des ordonnances hospitalières de 1996. Cette première section, en retraçant l'évolution des modes de financement hospitaliers en France, a pour objectif d'explicitier les relations stratégiques, qui existent aujourd'hui entre l'hôpital assurant le service public hospitalier et l'autorité publique.

1. Le prix de journée et la Dotation Globale de Fonctionnement

1.1. Le prix de journée

Introduit au XIX^{ème} siècle, ce prix couvre initialement les frais d'hôtellerie des malades, c'est-à-dire l'hébergement et la restauration. Progressivement, le coût du facteur travail, des médicaments, l'amortissement et la maintenance des investissements, les frais financiers entrent dans l'évaluation du prix de journée. Comme une facture est établie pour chaque patient sur la base de la durée d'hospitalisation, l'intensité des soins n'entre pas dans le calcul. Sans remettre en cause le prix de journée, les ordonnances du Professeur R. Debré de décembre 1958 ont pour objectif de faire de l'hôpital public le centre de référence de la médecine. Toutefois, l'essoufflement progressif de la croissance économique associé à la révolution technologique forte consommatrice de moyens financiers (imagerie, nouvelles molécules, etc...) fait prendre conscience que "*la santé n'a pas de prix mais un coût*" [Maesson, 1989 ; p. 26]. De fait, les réformes hospitalières du 31 décembre 1970 précisent davantage les missions, les moyens d'action et les structures du service public hospitalier. Elles stipulent d'ores et déjà qu'une réforme de tarification doit intervenir dans un délai d'un an après leur promulgation¹. Il faut néanmoins attendre 1983 pour que cette réforme intitulée « Dotation Globale de Financement (D.G.F.) » intervienne. Entre 1970 et 1983, les critiques à l'égard du prix de journée sont de plus en plus vives. Ce mode de financement, qualifié

¹ Article 52 de la loi hospitalière du 31 décembre 1970.

d'inflationniste, n'incite pas les hôpitaux à maîtriser leurs dépenses de santé. Le système *a priori*, encourage à prolonger les séjours des patients. Les statistiques ne traduisent, cependant, pas cette incitation. En effet, la Durée Moyenne de Séjour (D.M.S.) passe de 12,8 jours à 8,2 jours de 1974 à 1986. Les critiques mettent en exergue la rémunération unique malgré la forte hétérogénéité des activités médicales. Au tournant des années 1970-80, le prix de journée est devenu un mode de fonctionnement inadapté [Lacronique, 1981 ; Bordeloup, 1983]. C'est la raison pour laquelle la loi du 4 janvier 1978 autorise le gouvernement à expérimenter deux systèmes de tarification et de gestion des hôpitaux, d'une part le prix de journée éclaté, d'autre part le budget global. Le premier distingue le coût d'hébergement proportionnel à la durée du séjour, le coût des frais courants et d'administration sous forme forfaitaire, enfin les coûts individualisables à une technique de soins. Le second système prend la forme d'un budget global, qui couvre les charges d'exploitation prévisionnelles et évolue chaque année selon un taux directeur. C'est le second système de tarification, qui sera retenu et officialisé en 1983, compte tenu du succès des premières expériences [Renou, 1983].

1.2. La Dotation Globale de Fonctionnement

La réforme hospitalière du 11 août 1983, relative aux établissements publics et P.S.P.H., fixe les conditions d'élaboration des budgets, les nouvelles règles de la comptabilité hospitalière, les modalités et sources de financement [Clément, 1985]. La D.G.F. correspond à la différence entre la totalité des charges d'exploitation inscrites au budget général et la totalité des recettes correspondantes, autres que la dotation globale. La D.G.F. est mise en place progressivement de 1984 pour les centres hospitaliers régionaux à 1987 pour les services mobiles d'urgence. Cette D.G.F. s'accompagne d'une réforme de gestion [Fontaine, 1985]. Des objectifs sont élaborés par centre de responsabilité (gestion financière, gestion des patients, etc...) et des moyens sont alloués pour réaliser ces objectifs. Ainsi, des écarts susceptibles de provenir de la nature des contrôles peuvent être signalés entre les objectifs et les budgets. En effet, le contrôle de gestion est basé sur un coût - performance, tandis que le contrôle budgétaire est fondé sur les moyens consommés rapportés aux moyens alloués. Parallèlement, les centres d'activité deviennent des sous-ensembles homogènes des centres de responsabilité (qualité des soins, qualité de l'accueil, etc...)². La gestion, guidée par les objectifs, se déconcentre et chaque centre de responsabilité doit élaborer son TAbteau Prévisionnel

² Loi du 24 juillet 1987.

d'Activités et de Moyens (T.A.P.A.M.)³. Il sert de négociation, voire de contractualisation entre les responsables des centres de responsabilité et le directeur de l'hôpital. Cette nouvelle comptabilité analytique [Cibert, 1986] fournit une aide à la décision stratégique pour les directeurs d'hôpitaux, sert de base de comparaison inter-hospitalière et rend possible un calcul de coût par Groupes Homogènes de Malades (G.H.M.) [Hatchuel, 1983].

Mais la réforme hospitalière du 11 août 1983, qui visait notamment à maîtriser les dépenses hospitalières, à améliorer la gestion de l'hôpital, à accroître la responsabilité économique des gestionnaires et des médecins, connaît une critique croissante des professionnels de santé [Bandelier, 1983 ; Cadène, 1983]. En 1989, le Syndicat National des Cadres Hospitaliers (S.N.C.H.) refuse même le Budget Global [Vincent, 1989]. Ainsi, d'une expérimentation favorable au budget global, nous passons au refus de ce mode de financement lors de sa généralisation. Les principales explications sont la faiblesse des taux directeurs d'évolution annuelle en raison d'une conjoncture économique défavorable et de l'inflation, les retards importants pris dans l'application des réformes de gestion (qui dans la plupart des hôpitaux n'ont jamais été appliquées) et les rentes de situation que s'approprient les hôpitaux à faible productivité. Cette réforme est nuisible pour les hôpitaux dont l'activité augmente sensiblement car le montant de la D.G.F. n'est pas fonction de l'activité des établissements de santé, mais du taux directeur fixé par le Ministre de la Santé. Ce taux sert de référence pour la progression conseillée des dépenses d'hospitalisation. Une enveloppe départementale reste, toutefois, à disposition du Préfet pour compléter le cas échéant la D.G.F. En résumé, le mode de financement des hôpitaux publics et P.S.P.H. n'est pas satisfaisant, parce qu'il n'incite pas à la performance. En effet, la tutelle⁴ se réapproprie le plus souvent les résultats excédentaires qui pourraient servir à l'autofinancement des investissements, et la concurrence entre les établissements de santé n'existe pas [Majnoni d'Intignano, 1989]. Selon le Syndicat National des Cadres Hospitaliers (S.N.C.H.), le seul aspect positif de la dotation globale concerne l'amélioration de la trésorerie des hôpitaux, puisque chaque douzième de la dotation est versé mensuellement. Malgré les critiques, cette réforme n'est pas remise en cause par les gouvernements qui lui succèdent. Néanmoins, la D.G.F. est corrigée selon l'activité de l'hôpital évaluée par le P.M.S.I. à partir de 1995.

³ Décret du 11 août 1983.

⁴ Article 63 du décret du 11 août 1983.

2. Le Programme de Médicalisation du Système d'information dans le cadre de la Dotation Globale de Fonctionnement

2.1. Du Projet au Programme de Médicalisation du Système d'Information

Jean de Kervasdoué, directeur des hôpitaux en 1981, initialise le projet de médicalisation du système d'information afin de mesurer l'activité médicale en termes de produits [de Kervasdoué, 1981 ; Ministère de la Santé, 1981]. L'hôpital est alors conçu comme une entreprise à produits multiples [Girond, 1987] et l'approche synthétique des produits s'effectue à partir de la classification en G.H.M., directement inspirée des Diagnoses Related Groups (D.R.G.'s) nord-américains. Ces derniers, développés par l'équipe du Professeur R.F. FETTER [Fetter et *al.*, 1980]⁵, ont été expérimentés aux Etats-Unis en 1983, puis généralisés en 1987 afin de contrôler la croissance soutenue des dépenses de santé. Le projet français se transforme rapidement en un Programme de Médicalisation du Système d'Information (P.M.S.I.), qui permet alors de suivre l'évolution des case-mix dans chaque hôpital ; le case-mix correspondant à l'éventail des malades de courts séjours dans l'établissement de santé, autrement dit la répartition des malades de courts séjours d'un établissement par groupes homogènes. Aux Etats-Unis, ce système est utilisé pour la tarification des soins et si le coût de revient de la prise en charge d'un patient est supérieur au coût facturé, alors l'établissement de santé en supporte la perte financière. Dans le cas contraire, il conserve son profit. En France, il s'agit plus d'un outil de négociation et de comparaison dans le cadre du budget global. Il permet d'évaluer un coût de revient complet et favorise une meilleure connaissance de la production hospitalière. Le P.M.S.I. est donc un outil médico-économique conçu pour décrire l'activité des établissements de santé publics et privés [Ministère du Travail et des Affaires Sociales, 1996]. Ce dernier concerne les courts séjours de Médecine, Chirurgie, Obstétrique (M.C.O.) des établissements hospitaliers, sachant qu'un séjour correspond à la prise en charge par un hôpital d'un patient, pour une durée de quelques heures à plusieurs jours, sous la forme d'un épisode de soins continu. Les établissements de santé publics et privés de plus de cent lits M.C.O. en court séjour évaluent et analysent leur activité grâce à un recueil systématique et standardisé de données et aux

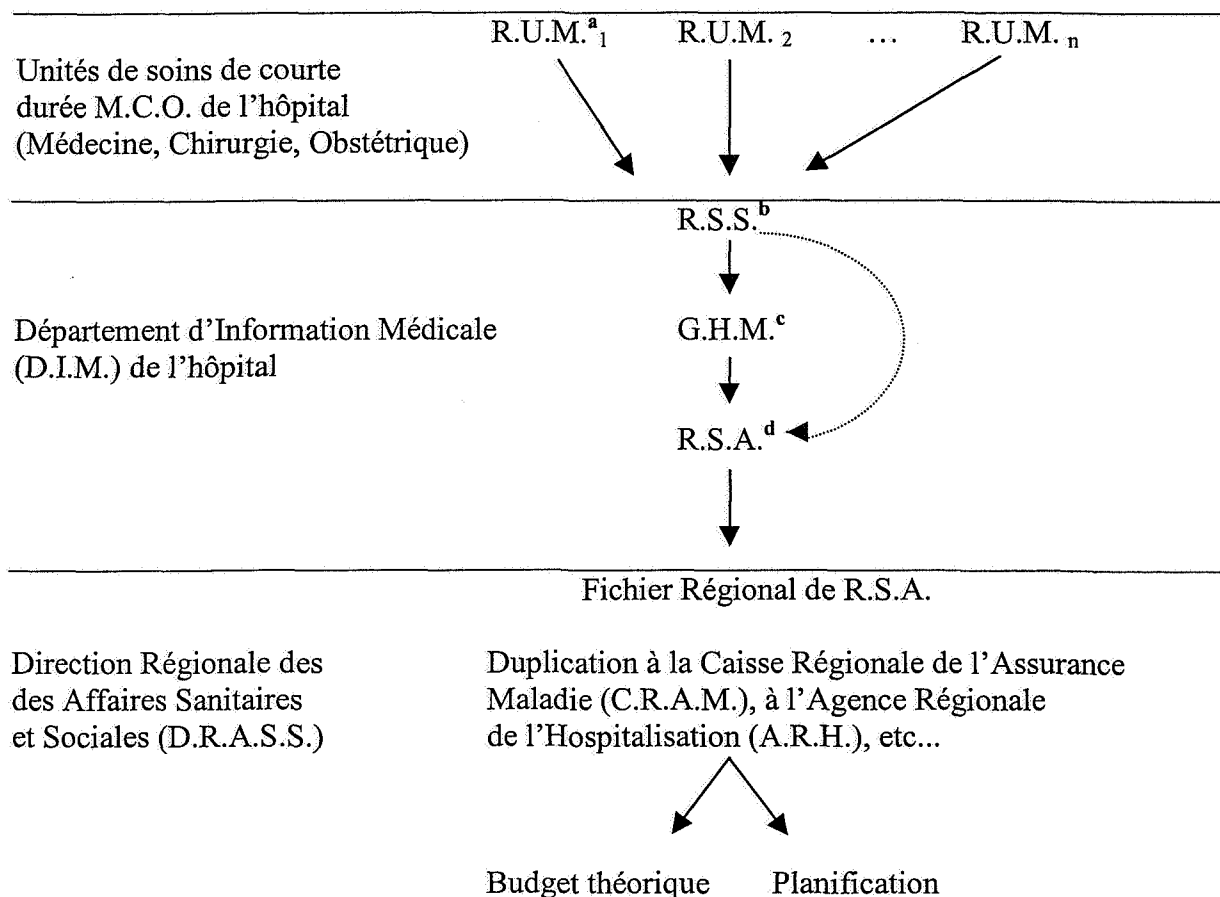
⁵ Les travaux du Professeur R.F. Fetter conduisent à la première classification DRG's, qui compte 470 DRG's en 1983 (499 en 2000). D'autres classifications expérimentales ont été créées aux Etats-Unis, comme le groupage « all-patient-DRG's » en 1988-90 ou encore « all-patient-refined DRG's » de 1991 [Hartmann et *al.*, 2001].

traitements automatisés de l'information médico-administrative⁶.

2.2. Les mécanismes du Programme de Médicalisation du Système d'Information

Avant de détailler le processus du P.M.S.I., nous proposons une vision globale de son fonctionnement à partir de la figure ci-dessous.

Figure 1
Production, traitement et utilisation de l'information médicale



^a R.U.M. : Résumé d'Unité Médicale ; ^b R.S.S. : Résumé de Sortie Standardisé ; ^c G.H.M. : Groupe Homogène de Malades ; ^d R.S.A. : Résumé de Sortie Anonyme.

Source : [Ministère du Travail et des Affaires Sociales, 1996 ; p.8].

Tout séjour hospitalier classé « court séjour M.C.O. » fait l'objet d'un Résumé de Sortie Standardisé (R.S.S.)⁷ constitué par un ou plusieurs Résumés d'Unités Médicales (R.U.M.). Chaque R.U.M., complété par le médecin responsable du patient, contient pour l'essentiel le numéro d'inscription au Fichier National des Etablissements Sanitaires et Sociaux

⁶ Loi 91-748 du 31 juillet 1991.

(F.I.N.E.S.S.), le numéro de R.S.S., la date de naissance, le sexe, le code postal, le numéro d'unité médicale, la date et le mode d'entrée dans l'unité, la date et le mode de sortie de l'unité, le type de mutation ou de transfert, le diagnostic principal, le ou les diagnostics associés. Le diagnostic, qui selon le médecin responsable du patient en fin de séjour a mobilisé l'essentiel de l'effort médical et soignant principal, devient le diagnostic principal du R.U.M. Les diagnostics associés concernent en général les affections et les complications associées au diagnostic principal. L'information contenue dans les R.U.M. est codée selon deux nomenclatures standardisées, d'une part la Classification Internationale des Maladies (C.I.M.), d'autre part le Catalogue des Actes Médicaux (C.d.A.M.). La C.I.M. permet de coder les diagnostics contenus dans les R.U.M., tandis que le C.d.A.M.⁸ publié pour la première fois en 1985 revêt deux objectifs. Il permet d'identifier et de coder les actes réalisés pendant le séjour du patient et de mesurer la consommation en ressources humaines et matérielles pour réaliser l'acte. Cette mesure repose sur les Indices de Coûts Relatifs (I.C.R.), l'Indice d'Activité Médicale (I.A.M.), l'Indice d'Activité Soignante (I.A.S.) et l'Indice de Consommation des Ressources Matérielles (I.C.R.M.). Ces deux nomenclatures sont régulièrement mises à jour par le Pôle d'Expertise et de Référence National des Nomenclatures de Santé (P.E.R.N.N.S.). Le C.d.A.M. se structure en cinq champs : le champ alpha concerne les actes de diagnostics et les actes thérapeutiques, le champ bêta les actes d'anesthésiologie, le champ gamma l'imagerie médicale, le champs mu la radiothérapie et le champ omega les actes de réanimation. Tout R.S.S. est classé dans un des G.H.M. Il faut souligner que le système des D.R.G.'s nord-américains est essentiellement adapté aux actes de chirurgie. Ceci explique que le P.M.S.I. apprécie parfaitement le coût de la chirurgie et moins bien les actes médicaux et obstétricaux. En 1995, le système français comptait 512 G.H.M., dont 462 G.H.M. d'une durée de séjour supérieure à 24 heures (hospitalisation complète), 46 G.H.M. d'une durée du séjour inférieure à 24 heures (hospitalisation en ambulatoire) et quatre G.H.M. relatifs aux erreurs et séjours inclassables. L'imputation d'un séjour dans un G.H.M. repose sur un algorithme de tri comprenant trois tests :

- (1) Le premier test, temporel, impute directement les séjours inférieurs à 24 heures dans la Catégorie Majeure (C.M.) numéro 24.
- (2) Les autres séjours font l'objet d'un second test sur le diagnostic principal et sont classés dans une des 23 Catégories Majeures de Diagnostic (C.M.D.). S'il s'agit d'un séjour mono-

⁷ Arrêté du 03 octobre 1985.

unité, le diagnostic principal du R.U.M. devient celui du R.S.S. Dans le cas contraire, le diagnostic principal est obtenu selon le schéma suivant : si un R.U.M. contient un acte classant opératoire, alors le diagnostic principal de ce R.U.M. devient celui du R.S.S. S'il n'existe pas d'acte opératoire classant ou si plusieurs R.U.M. en contiennent, alors le diagnostic principal du R.U.M. dont la durée de séjour est la plus longue devient le diagnostic principal du R.S.S. Si les séjours sont de durée équivalente, le diagnostic principal de la dernière unité médicale fréquentée par le patient devient le diagnostic principal du R.S.S.

(3) Un troisième test concerne la hiérarchie entre les actes classant propres à chaque C.M.D. Si un R.S.S. contient au moins un acte classant opératoire, le G.H.M. est chirurgical. Dans le cas contraire, le G.H.M. est médical. Les autres variables, qui interviennent dans l'algorithme de tri, sont les Complications et/ou Morbidité Associées (C.M.A.) et l'âge. Les R.S.S. sont ensuite transformés en Résumés de Sortie Anonymes (R.S.A.) et transmis à la Direction Régionale des Affaires Sanitaires et Sociales (D.R.A.S.S.), qui les transmet à son tour à l'A.R.H., ainsi qu'aux organismes de Sécurité Sociale. Les principales transformations pour passer des R.S.S. aux R.S.A. sont les suivantes : le numéro de R.S.S. disparaît, la date de naissance devient l'âge du patient en années (ou mois s'il s'agit d'un nourrisson), le code postal est transformé en code géographique, le nombre d'unités médicales fréquentées se substitue au numéro de l'unité ou des unités médicales, les dates d'entrées et de sorties deviennent des durées de séjours, le diagnostic principal par unité disparaît, seul le diagnostic principal du R.S.S. subsiste dans le R.S.A.

L'échelle nationale des coûts attribue à chaque G.H.M. un certain nombre de points I.S.A, qui permet d'évaluer l'activité produite. Cette évaluation est fondée sur un échantillon d'établissements hospitaliers publics et d'établissements privés participant au service public (P.S.P.H.). Elle intègre les dépenses directes affectées aux patients (matériel à usage unique, produits sanguins, médicaments coûteux, actes réalisés à l'extérieur, etc...), les actes exprimés en lettres clés et le nombre de journées d'hospitalisation au cours d'un séjour. De plus, l'ensemble des charges directes de fonctionnement est affecté à chaque service clinique ou médico-technique. Les activités de logistique et de structure sont réparties grâce à des unités d'œuvre. En résumé, le coût d'un séjour hospitalier comprend les dépenses directes, les dépenses des actes médico-techniques et des séjours dans les différentes unités médicales, les dépenses de restauration, de blanchisserie, de logistique et d'administration. En 1994, sur les

⁸ Le C.D.A.M. et la nomenclature générale des actes professionnels qui sert à la facturation des actes

43 établissements de soins volontaires pour servir de base à cette évaluation, seuls 22 établissements remplissaient les conditions nécessaires (comptabilité analytique suffisamment développée, etc...). Les coûts obtenus ont été redressés en pondérant les séjours en fonction de la sur ou sous-représentation de la catégorie d'établissements dont ils provenaient. Cette évaluation médico-économique a permis de dresser l'échelle nationale des coûts, construite à partir du coût médian observé sur l'échantillon. Par convention, le G.H.M. 540 intitulé « *Accouchement par voie basse sans complication* » génère 1000 points I.S.A., les autres G.H.M. se positionnant par rapport au G.H.M. 540.⁹ Le point I.S.A. s'avère une unité composite, qui permet de mesurer l'activité hospitalière. En 1995, l'étendue était de 89 points I.S.A. pour le G.H.M. 682 « *Séance de radiothérapie* » à 18657 points I.S.A. pour le G.H.M. 151 « *Transplantation cardiaque* ». Les fichiers de R.S.A. sont ensuite transmis à la D.R.A.S.S. au 30 septembre de l'année n pour les patients sortis au cours du premier semestre de l'année n et au 31 mars de l'année $n+1$ pour les patients sortis au cours du deuxième semestre de l'année n .¹⁰ Ces informations sont traitées au sein de la D.R.A.S.S. grâce au logiciel de Mesure de l'Activité HOSPitalière (M.A.HOS.). Chaque D.R.A.S.S. constitue annuellement une base régionale des R.S.A. valorisée en I.S.A., puis transmet ses données à d'autres autorités sanitaires régionales, ainsi qu'à la mission P.M.S.I. de la Direction des Hôpitaux, chargée d'élaborer la base nationale des R.S.A. Le P.M.S.I. permet d'une part à la tutelle d'évaluer annuellement la valeur régionale du point I.S.A. en francs en divisant au niveau régional les dépenses M.C.O. court séjour par les points I.S.A. générés par cette activité. Il permet, d'autre part, à l'hôpital i d'élaborer la valeur de son point I.S.A. noté $VISA_i$ ¹¹ et sa dotation théorique notée $DGFT_i$.¹² Cette dotation théorique de l'hôpital i est alors comparée à son budget réel noté DGF_i , c'est-à-dire la D.G.F. et les recettes hospitalières annexes. Cette comparaison donne lieu à un pourcentage de sur ou sous-dotation noté δ_i .¹³ Nous prenons une illustration simple en considérant l'hôpital A , qui produit 12500 points

fusionneront à partir de 2001 et deviendront le catalogue commun des actes médicaux.

⁹ Nombre de points I.S.A. affectés au G.H.M. _{i} = $\frac{\text{Coût médian G.H.M. } i}{\text{Coût médian G.H.M. 540}} \times 1000$

¹⁰ Article L.710-7 du Code de la Santé Publique devenu les articles L.6113-8 et L.6113-9 de la nouvelle partie législative du Code de la Santé Publique.

¹¹ La valeur du point I.S.A. de l'hôpital i , s'écrit $VISA_i = \frac{DGF_i}{ISA_i}$

¹² La dotation théorique de l'hôpital i , s'écrit $DGFT_i = \sum ISA_i \times \text{valeur du point I.S.A. régional}$.

I.S.A., auquel lui est rajouté la valorisation des consultations externes soit, par exemple, 1000 points I.S.A. Le nombre total de points I.S.A s'élève à 13500. En supposant que le point I.S.A. régional soit de 15 francs (2,3 €), la dotation théorique de cet hôpital $DGFT_A$ correspond à 13500×15 francs, soit 202500 francs (30870,9 €). Or, si le budget M.C.O. court séjour DGF_A versé par l'autorité publique à l'hôpital A s'élève à 133245 francs (20313,1 €), le pourcentage de sous-dotation δ_i s'établit à $-34,2\%$ soit $(((133245/202500)-1) \times 100)$. Cette comparaison permet, le cas échéant, de pénaliser les hôpitaux sur-dotés. En région Rhône-Alpes, par exemple, lorsque l'hôpital se trouve en sur-dotation au cours d'un exercice, l'A.R.H. prélève 5% de cette sur-dotation sur la D.G.F. de l'exercice suivant. Ces prélèvements atteignaient 39 millions de francs en 1999 [A.R.H.R.A., 2000]. Une certaine concurrence, qui se traduit par une course aux point I.S.A., apparaît alors entre les hôpitaux.

2.3. Les principales innovations de la classification en G.H.M. (version 6)

La sixième version de la classification en G.H.M., récemment mise en place [Ministère de l'emploi et de la solidarité, 2000], comprend 23 C.M.D., la C.M. 24 (séjours inférieurs à 24 heures) et la C.M. 27 (erreurs, séjours inclassables). Cinq innovations attirent particulièrement notre attention.

2.3.1. La création d'un diagnostic relié

Ce diagnostic, relié au diagnostic principal, permet d'effectuer un test supplémentaire lors de l'imputation d'un séjour dans un G.H.M. Ainsi, tout diagnostic susceptible d'éclairer le contexte pathologique doit figurer en diagnostic relié. Prenons deux exemples en cancérologie : un patient, atteint d'un cancer aux bronches, est hospitalisé pour une chimiothérapie. Le diagnostic principal correspond au code de la chimiothérapie (Z51.1), tandis que le diagnostic relié reste le cancer des bronches. Dans notre second exemple, les médecins découvrent, lors d'une surveillance d'un cancer du sein, une métastase osseuse. La métastase devient le diagnostic principal et le cancer du sein constitue le diagnostic relié.

2.3.2. La ventilation des diagnostics associés

Les diagnostics associés sont désormais ventilés en deux classes, d'une part le diagnostic associé significatif, d'autre part le diagnostic associé documentaire. Le diagnostic associé

¹³ Le coefficient de sur ou sous-dotation noté δ_i , s'écrit $\delta_i = \left(\frac{DGF_i}{DGFT_i} - 1 \right) \times 100$

significatif qualifie le séjour du patient et induit des effets sur le volume de soins consommé au cours du séjour à l'hôpital. Il s'agit de toute morbidité ayant donné lieu à une prise en charge diagnostique ou thérapeutique au cours du séjour dans le cadre d'une affection nouvelle ou de l'évolution d'une affection connue. Le diagnostic associé documentaire qualifie le patient. Il peut s'agir, par exemple, de résultats d'examens anormaux.

2.3.3. Les codes imprécis ne figurent plus en diagnostic principal

Les codes trop imprécis, comme le T14.2 « *Fracture d'une partie du corps non précisée* », sont interdits en diagnostic principal. Il est toujours possible, en effet, de consulter le dossier médical du patient, afin d'obtenir plus de détails.

2.3.4. La nouvelle valorisation de la radiothérapie

Au seul G.H.M. 682 de radiothérapie se substituent désormais cinq G.H.M., qui intègrent davantage l'hétérogénéité de la consommation des ressources en fonction du type de radiothérapie. Il s'agit des G.H.M. 685 « *Préparation à une irradiation externe avec une dosimétrie tridimensionnelle* », G.H.M. 686 « *Autres préparations à une irradiation externe* », G.H.M. 687 « *Techniques spéciales d'irradiation externe, en séances* », G.H.M. 688 « *Techniques complexes d'irradiation externes, en séances* », G.H.M. 689 « *Autres techniques d'irradiation externes, en séances* ». Chaque G.H.M. donne lieu à un nombre de points I.S.A. spécifiques. Un travail similaire est actuellement en cours pour améliorer la valorisation de la chimiothérapie.

2.3.5. La valorisation des soins palliatifs et des greffes de moelle

La valorisation des G.H.M. de soins palliatifs (G.H.M. 669 et 675) et du G.H.M. des greffes de moelle (G.H.M. 584) incorpore la durée de séjour. Au delà d'une certaine durée de séjour, une rémunération variable complète la rémunération fixe du G.H.M. Pour chacun de ces G.H.M., une borne inférieure et supérieure du temps de séjour, ainsi qu'une valorisation journalière en I.S.A. sont définies. Concernant le G.H.M. 584, la borne inférieure est de 46 jours, la borne supérieure de 133 jours et la valorisation journalière de 400 points I.S.A. Si un patient séjourne 63 jours, la valorisation de ce séjour en points I.S.A. atteint 18367 points I.S.A.¹⁴ ; au delà d'un séjour de 133 jours, la rémunération est plafonnée à 46367 I.S.A.¹⁵

¹⁴ Soit 11567 points I.S.A. au titre de la rémunération fixe du G.H.M. 584 et 6800 points I.S.A. au titre de la rémunération variable [(63-46) × 400].

2.4. Les limites du P.M.S.I.

Le P.M.S.I. présente des limites inhérentes à son fonctionnement, bien qu'il constitue une avancée majeure dans l'évaluation médico-économique des activités de soins et soit régulièrement amélioré.

2.4.1. Une comparaison inter-annuelle difficile

L'évolution des C.I.M., les suppressions et créations de G.H.M., les revalorisations en I.S.A. d'un même G.H.M. modifient les volumes des points I.S.A. d'une période à l'autre. La comparaison inter-annuelle de la valeur du point I.S.A. s'avère, dès lors, bien délicate. La seule issue consiste à reprendre l'ensemble de l'activité de l'année n et de la mesurer à travers la classification de l'année n-1. Le diagnostic principal du R.S.S. peut également masquer la réalité en terme d'effort médical. Une solution pourrait être de choisir systématiquement comme diagnostic principal du R.S.S., le diagnostic principal du R.U.M. qui maximise les points I.S.A. [Coca, 1998].

2.4.2. Le P.M.S.I. et les pratiques médicales : l'exemple de la prise en charge des cures de chimiothérapie

L'hôpital peut pratiquer une cure de chimiothérapie en hospitalisation complète ou en ambulatoire si les données cliniques du patient le permettent. S'il s'agit d'une hospitalisation en ambulatoire, le diagnostic principal impute le séjour dans le G.H.M. 681 « *Chimiothérapie pour tumeur* », qui génère 203 points I.S.A. La rémunération théorique s'élève pour l'hôpital à 2591 francs (395 €) par jour¹⁶, soit 15545,8 francs (2369,9 €) pour une cure de six jours. Lors d'une hospitalisation complète, le séjour est imputé par exemple dans le G.H.M. 585 « *Chimiothérapie pour leucémie aiguë* », qui génère 2517 points I.S.A. soit une rémunération théorique de 32116,9 francs (4896,2 €). Compte tenu de la différence des rémunérations théoriques, la valorisation des G.H.M. en I.S.A. peut tout à fait induire certaines pratiques médicales. Si l'établissement de soins pratique la chimiothérapie en ambulatoire pour le confort du patient (ce dernier se retrouve quotidiennement dans son environnement familial), son activité au regard des autorités sanitaires sera plus faible et, par conséquent, la valeur de son point I.S.A. plus élevée. Nous rappelons qu'une valeur trop élevée du point I.S.A. d'un hôpital peut être financièrement sanctionnée par la tutelle. Une expérimentation, qui repose sur 200 séjours de chimiothérapie par C.R.L.C.C., est en cours avec le concours de la Mission

¹⁵ Soit $11567 + [(133 - 46) \times 400]$.

P.M.S.I. et du P.E.R.N.N.S. La typologie proposée pour évaluer les chimiothérapies distinguerait les chimiothérapies adjuvantes sans maladie visible et dans le cadre de la prise en charge initiale de la maladie, les chimiothérapies pour récurrence, les chimiothérapies néo-adjuvantes ou d'induction thérapeutique, les chimiothérapies de première ligne métastatique et enfin les chimiothérapies de deuxième ligne métastatique ou plus. Cette mesure permettrait d'adapter davantage la rémunération théorique de l'hôpital en fonction de l'intensité des soins, qui varient sensiblement selon la localisation tumorale, le protocole, les antécédents du patient. De plus, la rémunération théorique, basée sur la C.I.M. 10, n'inclut pas la classification T.N.M.¹⁷ propre à la cancérologie.

2.4.3. Le P.M.S.I. et le financement de l'innovation : l'exemple des molécules onéreuses de chimiothérapie

Les hôpitaux soumis au P.M.S.I. doivent tenir le difficile pari de la performance au plan médical et d'un point I.S.A. proche de la moyenne régionale. Ainsi, un établissement de soins, qui utilise les dernières molécules de chimiothérapie ou les protocoles les plus modernes, peut se retrouver en position de sur-dotation, alors que les coûts plus élevés proviennent de la volonté d'être plus performant en terme de taux de survie. Le système P.M.S.I., basé sur les coûts médians, pénalise donc l'innovation technique et limite l'utilisation des molécules onéreuses, dès lors que leur coût excède la rémunération du G.H.M. [Philip et al., 1999a et 1999b]. Le prix moyen du mabthéra, par exemple, (un anticorps contre les lymphomes malins) est estimé à 11674 francs (1779,7 €) par jour, alors que la rémunération théorique du G.H.M. de chimiothérapie en ambulatoire est, rappelons-le, de 2590,3 francs (394,9 €). Pour financer ces dépenses d'acquisition de l'innovation, en particulier les molécules dites onéreuses, le système P.M.S.I. attribue à l'hôpital des points I.S.A. complémentaires. Si le point I.S.A. de l'hôpital est supérieur au point I.S.A. régional alors cette mesure permet effectivement de limiter la sur-dotation et, par conséquent, les éventuelles sanctions financières. En revanche, si le point I.S.A. de l'hôpital est inférieur au point I.S.A. régional, alors ces points I.S.A. supplémentaires équivalent à de la « fausse monnaie », puisqu'il n'y a pas de dotation supplémentaire. En outre, le financement des molécules onéreuses par des points I.S.A. supplémentaires pose trois interrogations :

(1) A partir de quel prix une molécule de chimiothérapie devient-elle onéreuse ?

¹⁶ Nous avons pris la valeur du point I.S.A. de la région Rhône-Alpes en 1999, soit 12,76 francs (1,95 €).

(2) Comment passer du prix libellé en francs à un nombre de points I.S.A. ?

Les pharmaciens des C.R.L.C.C. et de l'Assistance Publique des Hôpitaux de Paris (AP-HP) calculent un prix moyen journalier de ces molécules et proposent de considérer une molécule comme onéreuse, lorsque son prix dépasse mille francs par jour. La notion d'anticipation des nouveaux médicaments mis sur le marché doit être introduite au sein du P.M.S.I., qui doit évoluer d'un outil d'évaluation médico-économique favorisant le *statu quo* autour d'un comportement moyen vers un outil d'incitation conforme à une politique de santé publique. Entre 1994 et 1999 lorsque le coût de la consommation totale en produits pharmaceutiques augmentait de 20%, celui des médicaments anticancéreux grimpait de 230% [Nadal, 2001]. Toutefois, tous les médicaments dispensés par la pharmacie hospitalière ne relèvent pas de la D.G.F. En effet, les établissements publics de santé peuvent facturer aux caisses d'assurance maladie la prise en charge des médicaments qu'ils dispensent aux patients ambulatoires. Il s'agit du régime de la rétrocession (hors rétrocession, l'augmentation entre 1994 et 1999 est de 150%).

(3) Comment identifier les molécules onéreuses qui servent à traiter le cancer et les autres maladies comme par exemple l'interferon ?

2.4.4. Le P.M.S.I. et les surcoûts liés à la pédiatrie

Les surcoûts liés à la prise en charge d'un patient en cancérologie pédiatrique sont rarement pris en compte par le P.M.S.I. Dans le cadre d'un Prélèvement de Cellules Souches du Sang Périphérique (P.C.S.S.P.), le G.H.M. 827 « *Greffes d'organes : ambulatoire sans acte opératoire* » est valorisé 582 points I.S.A., quel que soit l'âge du patient. Or, un enfant requiert la présence d'un médecin hématologue tout au long du P.C.S.S.P., soit trois heures environ contre seulement vingt minutes pour un adulte en début de prélèvement [Perrier *et al.*, 2000]. De même, une irradiation corporelle totale (ou Total Body Irradiation (T.B.I.) en anglais) chez un enfant requiert un temps de radiothérapeute, de physicien et de manipulateur plus important. Pourtant, malgré l'amélioration significative de la valorisation des actes de radiothérapie au cours de l'exercice 2000¹⁸, la pédiatrie ne fait toujours pas l'objet d'une valorisation indépendante.

¹⁷ Le T.N.M. agrège trois mesures : la taille de la tumeur la plus grande (T), le nombre de ganglions envahis (N) et la présence ou non de métastases (M).

¹⁸ Se référer au paragraphe 2.3.4. « *La nouvelle valorisation de la radiothérapie* », p. 24.

2.4.5. Une étude dynamique d'un modèle régional.

Afin de bien cerner les limites du P.M.S.I. comme instrument de régulation, il est possible d'élaborer un modèle dynamique simple.

(1) Le modèle : On considère une région dotée de n hôpitaux publics et P.S.P.H. h_i , soumis au P.M.S.I. A la date t l'établissement de santé h_i engendre un certain volume de points I.S.A. noté $p(t,i)$ et reçoit une D.G.F. notée $d(t,i)$. L'A.R.H. calcule en t la valeur régionale du point I.S.A. , notée $vr(t)$. Celle-ci sert à déterminer la dotation théorique de h_i , notée $dt(t,i)$ et son coefficient de sur ou sous-dotation, noté $\delta(t,i)$. On a ainsi

$$vr(t) = \frac{\sum_{i=1}^n d(t,i)}{\sum_{i=1}^n p(t,i)} \quad (1)$$

$$dt(t,i) = p(t,i) \times vr(t,i) \quad \forall i \in (1, n) \quad (2)$$

$$\delta(t,i) = \frac{d(t,i)}{dt(t,i)} \quad \forall i \in (1, n) \quad (3)$$

Alors

si $\delta(t,i) > 1$ l'élément h_i est en sur - dotation

si $\delta(t,i) < 1$ l'élément h_i est en sous - dotation

Si h_i est sur-doté, on note

$$c(t,i) = d(t,i) - dt(t,i) \quad (4)$$

et la somme de ces composantes constitue une « cagnotte », qui sera redistribuée à la période ultérieure aux établissements sous dotés, selon une certaine proportion (égale à 5% en Rhône-Alpes). La « cagnotte » redistribuée est ainsi égale à

$$ca(t,i) = 0,05 \times \sum c(t,i) \quad (5)$$

On suppose que la répartition entre les établissements de santé sous-dotés se fait à proportion de leur sous-dotation . C'est-à-dire qu'une sous-dotation est notée

$$sd(t,i) = dt(t,i) - d(t,i) \quad (6)$$

et on calcule des coefficients de répartition $\beta(t,i)$ tels que

$$\beta(t,i) = \frac{sd(t,i)}{\sum sd(t,i)} \quad (7)$$

Cela étant, les D.G.F. sont censées évoluer selon un coefficient directeur (pris ici égal à 1,5%) et le passage entre une période et la suivante est décrit par deux équations selon que l'hôpital h_i est sous ou sur-doté. On a respectivement

$$\begin{aligned} d(t,i) &= 1,015 \times d(t-1,i) + \beta(t,i) \times ca(t,i) & h_i \text{ sous - doté} \\ d(t,i) &= 1,015 \times d(t-1,i) - 0,05 \times c(t,i) & h_i \text{ sur - doté} \end{aligned} \quad (8)$$

(2) Le passage au modèle de simulation : La partie « motrice » du modèle est donnée par les deux équations structurelles de (8). On a des équations aux différences qui, compte tenu de l'ensemble des autres équations, se révèlent fortement non linéaires. Il n'est pas possible d'obtenir une solution analytique du modèle, même avec un nombre restreint d'établissements de santé. Il faut donc recourir à la simulation informatique.

Le modèle de simulation est développé sous Matlab ®. On le nomme « PISA » pour « Points I.S.A. ». Nous fournissons en annexe l'ensemble du modèle¹⁹. Sans perte de généralité et pour plus de maniabilité nous poserons que $n=4$. De fait on prend

$h_1 = \text{HCL}$	$h_2 = \text{CLB}$	$h_3 = \text{Firminy}$	$h_4 = \text{Crest}$
--------------------	--------------------	------------------------	----------------------

Avec H.C.L. = Hospices Civils de Lyon

Le principal problème est d'engendrer des données de départ en adéquation avec les données observées. Fondamentalement, nous partons de ce que l'on appellera l'activité d'un hôpital h_i , notée $A(t,i)$. Celle-ci est assimilée au nombre de R.S.A. Les données dont nous disposons pour les années 1996, 1997 et 1999 sont présentées dans le tableau suivant.

Tableau I
Nombre de R.S.A. pour quatre hôpitaux de la région Rhône-Alpes

RSA	HCL	CLB	Firminy	Crest
1996	233 305	16 011	11 595	2 396
1997	257 197	16 043	12 616	2 325
1999	263 635	18 138	12 714	2 493
Moyenne	251 379	16 731	12 308	2 405

Source : D.R.A.S.S. Rhône-Alpes

¹⁹ Se référer à l'annexe « Détail du modèle « PISA » », p. 214.

Dans la mesure où l'activité est censée varier, nous supposons qu'elle obéit à un processus gaussien et qu'à chaque période

$$A(t,i) \approx N(m_A, \sigma_A)$$

Pour ce qui est des moyennes, ce sont celles calculées dans le tableau antérieur. Elles sont données dans le vecteur ma du programme. Pour les écart-types, nous tablons sur le fait que, dans le cadre d'un processus aléatoire gaussien $N(m, \sigma)$, on a 99,994% des observations contenues dans un intervalle de la forme $m \pm 4\sigma$, avec σ choisis de façon adéquate pour qu'il n'y ait pas de valeurs aberrantes. Les écart-types sont fournis dans le vecteur $sigmaa$ du programme qui est tel que $sigmaa = ma/80$.

Pour passer de l'activité aux points I.S.A. nous utilisons un coefficient g qui n'est autre, initialement, que la moyenne des rapports observés

$$\frac{[\text{Total ISA}](i, t)}{[\text{Nb de RSA}](i, t)}$$

Les valeurs numériques observées pour les trois années considérées sont, pour ce qui du total des points ISA

Tableau II
Nombre d'I.S.A. pour quatre hôpitaux de la région Rhône-Alpes

Nb ISA	HCL	CLB	Firminy	Crest
1996	248 044 220	15 424 746	11 404 606	2 864 587
1997	335 500 073	19 220 816	14 749 473	3 138 723
1999	332 913 648	20 627 337	14 317 662	3 120 954
Moyenne	305 485 980	18 424 300	13 490 580	3 041 421

Source : D.R.A.S.S. Rhône-Alpes

A partir du nombre de R.S.A. et de points I.S.A. de chacun des établissements de santé, nous calculons les coefficients moyens g de départ

Tableau III
Coefficients moyens g de départ pour les quatre hôpitaux

g	HCL	CLB	Firminy	Crest
1996	1063,18	963,38	983,58	1195,57
1997	1304,45	1198,08	1169,11	1349,99
1999	1262,78	1137,24	1126,13	1251,89
Moyenne	1210	1100	1093	1266

Ces coefficients, tout comme l'activité, sont censés varier en suivant un processus aléatoire normal. Aussi on pose que

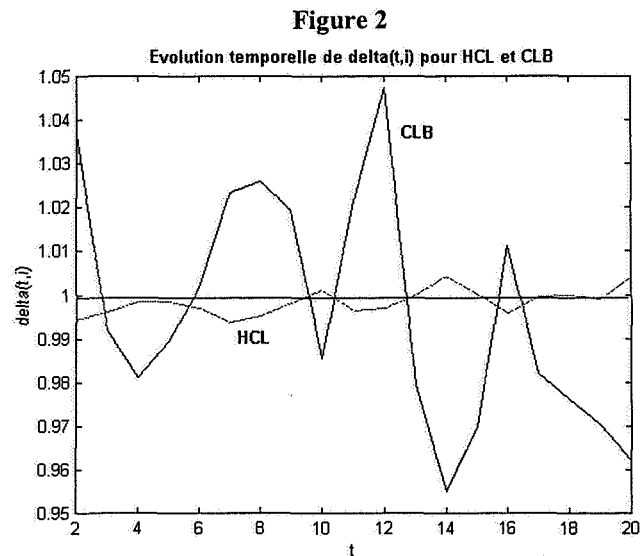
$$g(t,i) = N(m_g, \sigma_g)$$

Les moyennes sont celles données dans le tableau antérieur. Pour ce qui est des écart-types nous les avons pris égaux au 1/8 de l'étendue des valeurs observées de g. Soit le vecteur (30,29,23,19).

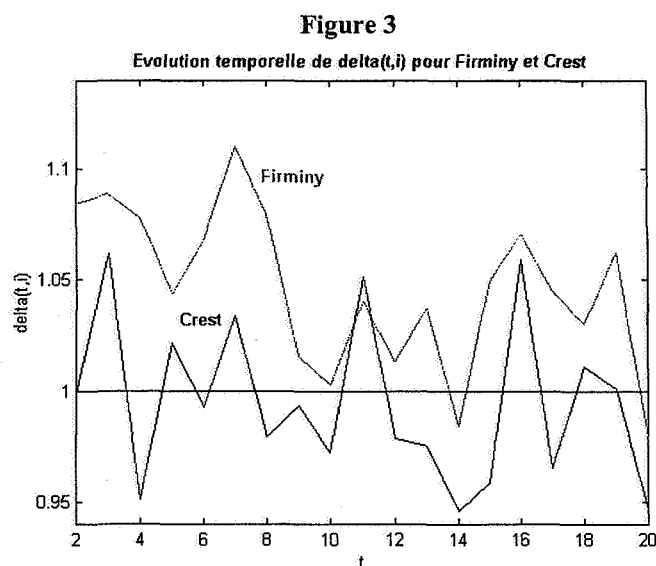
Les tirages aléatoires de nombres gaussiens se font sous Matlab à l'aide de la fonction *normrnd*. Les valeurs initiales $A(1,i)$, $g(1,i)$ et $p(1,i)$, $i \in (1,4)$, sont calculées dans une première boucle, hors de la boucle principale de calcul.

Pour ce qui est de la dotation initiale de chacun des établissements de santé h_i , compte tenu du rapport observé entre dotation financière et nombre total de points ISA au niveau de la région Rhône-Alpes, de l'ordre de 10, nous avons supposé que celle-ci était dans chaque cas issu d'un tirage gaussien de valeur moyenne $p(1,i)$ d'écart-type égal à $p(1,i)/80$. Le nombre aléatoire tiré étant multiplié par 10.

(3) Les principaux résultats : La période de temps étant l'année, nous avons réalisé des simulations sur 20 ans. De façon à voir si le mécanisme de « ponction - redistribution » fonctionne bien, il semble que le meilleur indicateur soit donné par l'évolution temporelle du coefficient de sur ou sous dotation. Pour les deux éléments les plus importants, nous obtenons le graphe suivant.

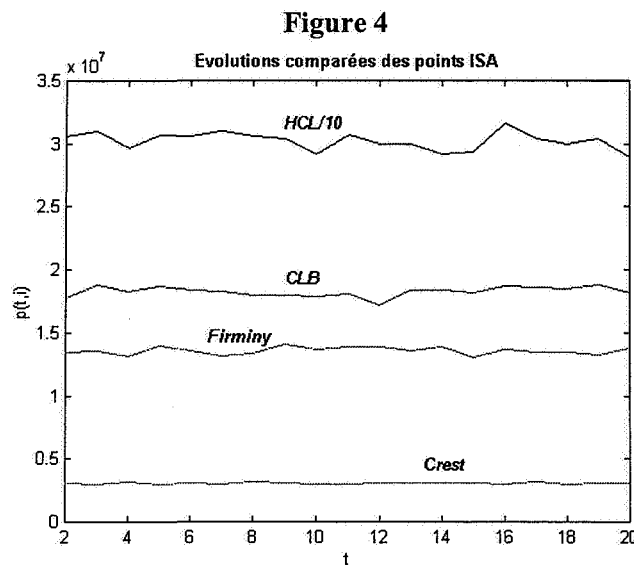


Les deux hôpitaux fluctuent autour de l'équilibre ($\delta(t,i) = 1$), la fluctuation des H.C.L. étant nettement moins prononcée que celle du C.L.B. Les H.C.L. fluctuent entre 0,99364 et 1,0042 alors que CLB fluctue entre 0,95506 et 1,0475. Les valeurs moyennes enregistrées sur la période sont respectivement de 0,99829 et 0,99630. On notera qu'elles sont légèrement inférieures à la valeur unité. Dans un cas comme dans l'autre, les fluctuations semblent se compenser. En outre, il semblerait que pour les H.C.L., on ait une légère tendance à la hausse avec l'inverse pour CLB. Sur un plan purement statistique on ne peut conclure, les ajustements linéaires par rapport au temps n'ayant pas des R^2 suffisants dans les deux cas. Pour ce qui est de Firminy et Crest, on a de même



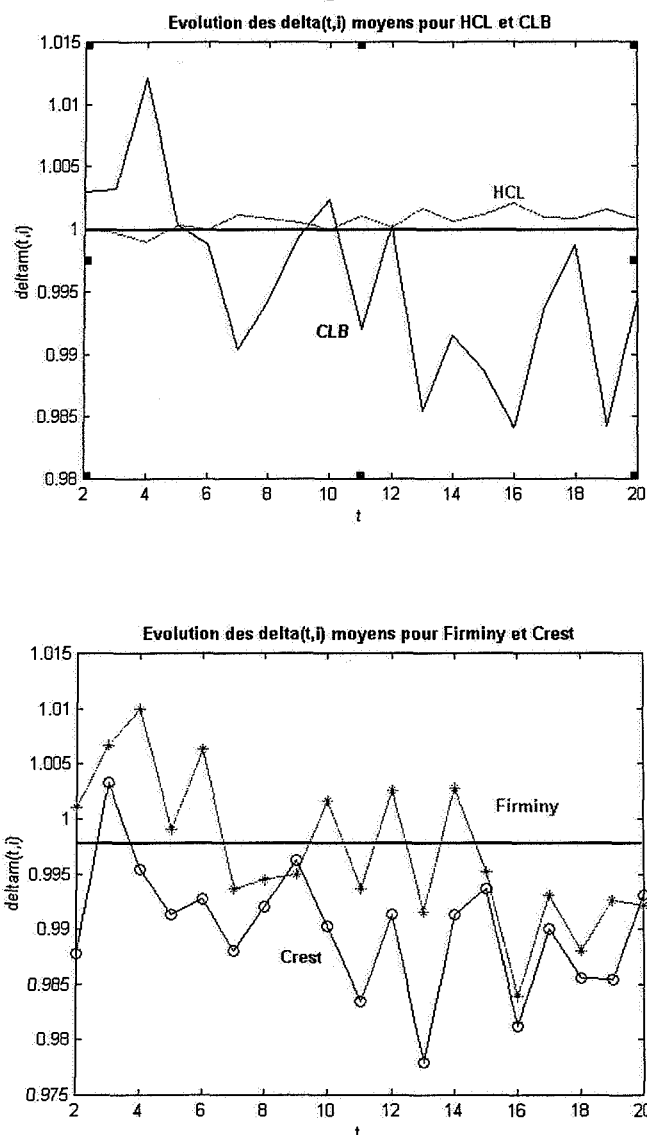
Dans ce cas, on a des évolutions quelque peu similaires avec cependant un décalage de valeurs, Firminy oscille entre 0,97940 et 1,1099 (avec une valeur moyenne de 1,0465 > 1), alors que Crest oscille entre 0,94576 et 1,0619 (avec une moyenne égale à 0,99464 < 1).

Si l'on veut tracer les évolutions comparées des points I.S.A. des quatre éléments, il est nécessaire de se livrer à un changement d'échelle pour les H.C.L., compte tenu de la disparité des valeurs. La figure correspondante est donnée ci-après.



Dans la mesure où nous utilisons des variables aléatoires, on pourrait arguer du fait que les tendances enregistrées ne sont qu'une possibilité parmi bien d'autres. Cela étant nous avons étendu notre modèle en répétant la simulation un certain nombre de fois ($m=50$). Les données pour les $\delta(t,i)$ sont alors stockées dans des tableaux multidimensionnels de la forme $\delta(t,i,k)$. Le format étant (20,4,50). Puis, en réalisant des moyennes sur la troisième dimension on obtient pour chaque hôpital h_i une tendance moyenne du coefficient de sur ou sous-dotatation $\delta_m(t,i)$. Nous donnons, ci-après, deux graphes d'évolutions temporelles comparées, qui parlent d'eux-mêmes.

Figure 5



Si on enregistre une moindre dispersion autour de la valeur d'équilibre (= 1), les tendances précédemment perçues se confirment. **En conclusion, le modèle dynamique montre que les H.C.L., qui génèrent pratiquement 30% des points I.S.A. de la région Rhône-Alpes, sont dans la quasi impossibilité de revenir à l'équilibre, c'est-à-dire à la valeur du point I.S.A. régional.**

Après avoir développé le fonctionnement du P.M.S.I, nous analysons comment sa mise en place s'est réalisée dans un contexte de réformes profondes des relations entre l'établissement de santé et sa tutelle. Ces dernières, dans le cadre des ordonnances Juppé de 1996, se sont nettement contractualisées. Nous consacrons le paragraphe suivant aux ordonnances de 1996.

3. Les ordonnances hospitalières de 1996

Parmi les trois ordonnances de 1996²⁰, l'ordonnance 96-346²¹ dont la création des agences régionales de l'hospitalisation est un des pivots [Caillet et *al.*, 1997], s'avère essentielle pour ce travail.

3.1. L'Agence Régionale de l'Hospitalisation

3.1.1. Les missions de l'A.R.H.

Les Agences Régionales de l'Hospitalisation (A.R.H.) créées dans chaque région et dans la collectivité territoriale de Corse ont pour mission de définir et de mettre en œuvre la politique régionale d'offre de soins hospitaliers, d'analyser et de coordonner l'activité des établissements de santé publics et privés et de déterminer leurs ressources.²² Chaque A.R.H. reçoit une part de l'enveloppe budgétaire nationale préalablement votée par l'Assemblée Nationale dans le cadre de la loi de financement de la Sécurité Sociale. Cette enveloppe, qui porte le nom d'Objectif National de Dépenses de l'Assurance Maladie (O.N.D.A.M.), s'élevait à 658,3 milliards de francs (100,36 milliards d'euros) en 2000, dont 255,7 milliards de francs (39 milliards d'euros) attribués au financement des établissements de santé sous D.G.F. [Ministère de l'Emploi et de la Solidarité, 2001a]. Seule la D.G.F. hospitalière fait partie de l'O.N.D.A.M., les produits de l'activité hospitalière (tickets modérateurs, forfaits journaliers, patients admis à titre payant, etc...) et les recettes subsidiaires (bénéfices des accords de recherche, etc...) y étant exclus. La répartition de l'O.N.D.A.M. entre les régions incombe pleinement au Ministre chargé de la Santé et de la Sécurité Sociale. Ce dernier établit, sur la base du P.M.S.I., une péréquation qui favorise les régions sous dotées au détriment des régions surdotées. Puis, il délègue aux A.R.H. la ventilation des enveloppes régionales entre les établissements de santé, sous contrainte du taux directeur.²³ La délégation attribuée aux A.R.H. par les Ministres chargés de la Santé et de la Sécurité Sociale est étendue.²⁴ Cette délégation est non seulement une délégation de contrôle mais aussi une délégation de pouvoir comme le montrent les développements ultérieurs.

²⁰ Ordonnances 96-344, 96-345 et 96-346 du 24 avril 1996.

²¹ Publiée au Journal Officiel de la République Française, n°98 du 25 avril 1996, p. 6324-6337.

²² Article L.710-17 de l'ordonnance 96-346 portant réforme de l'hospitalisation publique et privée devenu l'article L.6115-2 de la nouvelle partie législative du Code de la santé Publique.

²³ Se référer au paragraphe 1.2. « *La Dotation Globale de Fonctionnement* », p. 16.

3.1.2. Le directeur de l'A.R.H., Préfet sanitaire de région

Les A.R.H. sont administrées par une commission exécutive et dirigées par un directeur nommé en Conseil des Ministres, ce mode de nomination ayant des conséquences en termes de relations d'agence. La commission exécutive est présidée par le directeur de l'A.R.H. et regroupe à parité les représentants de l'Etat (D.R.A.S.S., etc...) et de l'assurance maladie (C.R.A.M., etc...). L'assurance maladie demeure cependant l'interlocuteur habituel des établissements privés à but lucratif et de la médecine de ville et les D.D.A.S.S. l'interlocuteur des établissements publics et des P.S.P.H. de son département. Leurs directeurs bénéficient d'une large délégation de signature du directeur de l'A.R.H. Certains domaines, comme la saisine de la chambre régionale des comptes, l'approbation des projets d'établissements et des contrats d'objectifs et de moyens, ne font pas contre l'objet d'aucune délégation de signature. La création de l'A.R.H. favorise le passage d'une étude séparée des dossiers par les services de l'Etat et par l'assurance maladie à une analyse partagée avec des échanges d'information. A l'ancienne vision des tutelles se substitue alors la position commune de l'A.R.H. Même si la C.R.A.M., les D.D.A.S.S., la D.R.A.S.S. et d'autres institutions conservent un rôle important, le directeur de l'A.R.H. est, en définitif, le véritable Préfet de sa région sur le plan sanitaire, c'est-à-dire le représentant de l'Etat. D'ailleurs, toute modification d'activité au sein d'un établissement de soins fait l'objet de son autorisation. La création, l'extension ou la conversion de tout établissement de santé public ou privé est soumise à l'autorisation du Ministre chargé de la santé et du directeur de l'A.R.H. En vue d'adapter le système hospitalier aux besoins de la population, le directeur de l'A.R.H. peut aussi demander à deux ou plusieurs établissements publics de santé de conclure une convention de coopération, de créer un syndicat inter-hospitalier ou un groupement d'intérêt public, voire de prendre une délibération pour créer un nouvel établissement public de santé par fusion des établissements concernés. Ce pouvoir engendre inéluctablement des comportements stratégiques.

3.1.3. L'Agence Régionale de l'Hospitalisation Rhône Alpes (A.R.H.R.A.)

Comme nous venons de le voir, une des missions de l'A.R.H.R.A. est d'allouer les ressources du budget de la nation destinées à la région Rhône-Alpes pour son secteur sanitaire. En 1997, la dotation budgétaire de la région Rhône-Alpes pour les établissements de santé publics et P.S.P.H. atteint 23,7 milliards de francs (3,61 milliards d'euros) [A.R.H.R.A., 1998a], soit

²⁴ Article L.710-18 du code de la santé publique devenu l'article L.6115-1 de la nouvelle partie législative du Code de la Santé Publique.

une augmentation de 0,7% par rapport à 1996. Par comparaison, les dépenses d'hospitalisation privée sont de l'ordre de 4 milliards de francs (0,61 milliards d'euros) auxquelles il faut rajouter 1,2 milliards d'honoraires (0,18 milliards d'euros). L'A.R.H.R.A. a utilisé une part de son budget pour financer la mise en œuvre du Schéma Régional d'Organisation Sanitaire (S.R.O.S.) et les priorités de santé publique définies pour l'année 1997 en l'occurrence les services d'urgence, les infections nosocomiales et la sécurité anesthésique. Pour l'année 1998, le budget de Rhône-Alpes progresse de 1,76 % pour s'établir à 24,2 milliards de francs (3,69 milliards d'euros) [A.R.H.R.A., 1998b]. Pour l'année 1999, la dotation de 24,8 milliards de francs (3,78 milliards d'euros) progresse de 2,09 % par rapport à 1998 [A.R.H.R.A., 1999a]. Alors qu'en 1998 la progression du budget de la région était nettement supérieure à la moyenne nationale (1,76 % contre 1,41 %), la dotation 1999 est très proche de la moyenne nationale (2,09 % contre 2,1 %). La région ne bénéficie donc plus de la péréquation interrégionale, dont les taux sont fixés par le Ministre de la Santé. Tous critères confondus et par habitant, elle se situe au 4^{ème} rang des régions les mieux dotées derrière l'Ile-de-France, la Corse et la région Provence-Alpes-Cote-d'Azur. Pour l'année 2000, la dotation budgétaire atteint 25,38 milliards de francs (3,87 milliards d'euros), soit une progression de 2,1 % contre 2,2 % en moyenne nationale [A.R.H.R.A., 2000].

3.2. Les fondements du contrat d'objectifs et de moyens

Les contrats d'objectifs et de moyens, institués par les ordonnances de 1996²⁵ sont issus de la confrontation entre le S.R.O.S. et les objectifs de chaque établissement de santé exprimés dans son projet d'établissement.

3.2.1. Le Schéma Régional d'Organisation Sanitaire et Sociale

Le Schéma Régional d'Organisation Sanitaire et Sociale (S.R.O.S.)²⁶ définit la répartition géographique des installations et des activités de soins. Il fixe pour une durée de cinq ans les priorités et les objectifs pour l'organisation sanitaire de la région. Tous les S.R.O.S. ont fait l'objet d'une révision en 1998, donnant lieu à un S.R.O.S. de deuxième génération sur la période 1999-2004. En Rhône-Alpes, la mise en place du S.R.O.S. (1999-2004) a été pilotée par secteur sanitaire, les secteurs étant définis par la carte sanitaire. Pour les programmes d'actions par secteur sanitaire, le S.R.O.S. donne priorité, par exemple, à la mise en place des réseaux de maternités et de cancérologie, à l'évolution de l'offre de soins de suite et de

²⁵ Circulaire DH/EO n° 97-22 du 13/01/97 relative aux contrats d'objectifs et de moyens.

rééducation. Concernant le développement d'actions partenariales d'intérêt régional, il prône notamment une commission des innovations diagnostics et thérapeutiques, un observatoire régional des plaintes, un serveur régional sur la gestion du risque sanitaire [A.R.H.R.A., 1999b].

La carte sanitaire, qui découpe la région en secteurs sanitaires d'au moins 200000 habitants, s'impose à tous les établissements de soins publics et privés, qu'ils participent ou non au service public. Elle est élaborée tous les cinq ans et arrête au sein de chaque secteur notamment les besoins pour les disciplines de M.C.O. [A.R.H.R.A., 1997].

Initialement signés par le Préfet de région, le S.R.O.S. et la Carte sanitaire relèvent depuis les ordonnances de 1996 du directeur de l'A.R.H. En Rhône-Alpes, il existe 11 secteurs sanitaires (Bourg-en-Bresse, Valence-Privas, Montélimar-Aubenas, Grenoble, Vienne-Bourgoin, Saint-Etienne, Roanne-Feurs, Lyon-Givors, Villefranche, Chambéry et Annecy) [A.R.H.R.A., 1997].

3.2.2. Le Projet d'Etablissement

Le projet d'établissement définit les objectifs généraux de l'établissement dans le domaine médical, des soins infirmiers et de la recherche bio-médicale mais aussi en ce qui concerne la politique sociale, le plan de formation, l'investissement pluriannuel et le système d'information. Le projet d'établissement intègre le projet médical. L'A.R.H., représentée par son directeur, approuve ou rejette les projets en fonction des orientations régionales définies par le S.R.O.S. et la carte sanitaire. Si l'A.R.H. approuve les projets, alors une négociation s'engage entre les deux parties pour définir les moyens essentiellement financiers nécessaires et parvenir aux objectifs fixés. Cette négociation aboutit alors dans un délai maximum de six mois au contrat d'objectifs et de moyens. Si l'A.R.H. rejette le projet, alors il doit être révisé puis proposé une seconde fois. Pour les Hospices Civils de Lyon (H.C.L.) par exemple, le projet d'établissement de deuxième génération (1999-2003) a été approuvé à l'unanimité par la commission médicale d'établissement et à une très large majorité par le conseil d'administration du 16 octobre 1998 [H.C.L., 2000]. Toutefois, en mars 1999, le directeur de l'A.R.H. refuse de l'approuver. Après la création d'une commission de concertation entre les H.C.L. et l'A.R.H., un *addendum* au projet est à nouveau proposé au directeur de l'A.R.H. Le coût du projet d'établissement est évalué à 2,6 milliards de francs (0,4 milliards d'euros). Après l'approbation du projet, la négociation devra aboutir au contrat d'objectifs et de

²⁶ Loi du 31 juillet 1991.

moyens.

3.3. Le contrat d'objectifs et de moyens

Le contrat d'objectifs et de moyens signé entre l'A.R.H. et chaque établissement de soins publics, P.S.P.H. ou privés à but lucratif résulte d'un compromis entre l'intérêt de santé publique exprimé dans le S.R.O.S. et la carte sanitaire, et les intérêts de l'hôpital exprimés dans son projet d'établissement. Il existe, par conséquent, depuis les ordonnances de 1996 une dimension régionale de la contractualisation.²⁷ Avant les ordonnances de 1996, la signature de contrats d'objectifs entre un établissement de soins, l'État, la Sécurité Sociale et les collectivités territoriales était facultative. Compte tenu du succès des quelques expériences, les ordonnances de 1996 astreignent à la signature de contrats d'objectifs et de moyens. Ces contrats, signés pour une durée de 3 à 5 ans, constituent un instrument de pilotage pour l'hôpital dans la mesure où ils intègrent des paramètres organisationnels, sanitaires, humains et financiers. Cet outil juridique et financier se veut souple, transparent et responsabilisant. Même s'il existe bien deux types de contrats d'objectifs et de moyens, selon qu'il s'agisse d'établissements relevant du régime de la dotation globale (public et P.S.P.H.) ou d'établissements sous contrat tripartite national de l'hospitalisation privée, leurs points communs sont nombreux [de Calan, 2000]. Nous détaillons, ci-dessous, quelques éléments communs aux contrats des établissements publics et privés (P.S.P.H. ou à but lucratif). Les contrats décrivent les transformations, que l'établissement s'engage à réaliser dans ses activités, son organisation, sa gestion, ses coopérations. Il définit aussi des objectifs en matière de qualité et de sécurité des soins, comme les délais de mise en œuvre de la procédure d'accréditation. Le contrat précise également les critères d'évolution des budgets des établissements de santé selon le degré de réalisation des objectifs. Il existe donc un dispositif de suivi du contrat d'objectifs et de moyens et un calendrier d'exécution. Un rapport annuel d'étape, ainsi qu'un rapport final sont fournis par l'hôpital à l'A.R.H. Des indicateurs ont été préalablement définis pour mesurer l'écart entre la situation initiale et celle constatée au terme du contrat. Si le contrat n'est pas exécuté, le directeur de l'A.R.H. est à même de mettre en œuvre les sanctions prévues à cet effet. Ce dernier peut même être résilié avant son terme par l'A.R.H., si l'établissement de santé manque gravement à ses obligations législatives, réglementaires ou contractuelles. Dès lors, la contractualisation devient un élément central de la vie des établissements de santé publics et privés. L'information délivrée par l'hôpital à

²⁷ Circulaire DH/EO n° 97-22 du 13 janvier 1997 relative aux contrats d'objectifs et de moyens.

l'A.R.H. et les négociations concernant d'éventuels surcoûts jouent alors un rôle stratégique pour justifier, par exemple, d'un point I.S.A. plus élevé que le point I.S.A. régional.

Les mécanismes du P.M.S.I., insérés dans le cadre de la contractualisation et des négociations des ordonnances de 1996, offrent une perspective intéressante de relation d'agence entre les autorités sanitaires et les établissements de soins.

3.4. Les contrôles de l'Agence Régionale de l'Hospitalisation et des hôpitaux

3.4.1. Les contrôles de l'Agence Régionale de l'Hospitalisation par l'Etat

Afin de prévenir d'éventuelles collusions entre l'A.R.H. et les établissements de soins, le Ministre de la Santé se dote de moyens législatifs pour la contrôler.²⁸ Ainsi, les A.R.H. sont soumises au contrôle économique et financier de la cour des comptes et au contrôle de l'inspection générale des affaires sociales.

3.4.2. Les contrôles des hôpitaux par l'Agence Régionale de l'Hospitalisation

Les contrôles des équipes de l'A.R.H. dans les établissements de soins prennent la forme d'audits complets d'un hôpital, d'enquêtes ponctuelles sur une activité donnée, de contrôles au titre de la sécurité sanitaire voire d'expertises médicales inter-établissements. Par exemple, les médecins membres de l'inspection générale des affaires sociales, les médecins inspecteurs de la santé publique et les médecins conseil des organismes d'assurance maladie ont accès aux dossiers médicaux des patients.

3.5. Les missions de l'Agence Nationale d'Accréditation et d'Évaluation en Santé

L'évaluation de la qualité des soins apparaît aux Etats-Unis au lendemain de la première guerre mondiale. Mais, il faut attendre les années 1950 et la création de la Joint Commission of Health Organization pour que l'évaluation de la qualité connaisse un développement important. Ce concept, qui apparaît en France au milieu des années 1970, intègre le droit hospitalier français en 1991, lors de la création de l'Agence Nationale pour le Développement de l'Évaluation Médicale (A.N.D.E.M.).²⁹ Dans le cadre des ordonnances de 1996, cette institution est remplacée par l'Agence Nationale d'Accréditation et d'Évaluation en Santé (A.N.A.E.S.). Disposant d'une mission plus large, l'A.N.A.E.S. élabore, valide et diffuse les

²⁸ Article L.710-17 du code de la santé publique devenu l'article L.6115-2 de la nouvelle partie législative du Code de la Santé Publique.

méthodes et les référentiels d'accréditation. Puis, elle accrédite les hôpitaux, c'est-à-dire porte une appréciation indépendante sur la qualité d'un établissement à l'aide des référentiels préalablement définis. La procédure d'accréditation s'impose à tous les établissements de santé publics et privés, P.S.P.H. ou non. Si dans un délai de cinq ans à compter de la publication de l'ordonnance 96-346 un établissement de santé n'a pas engagé cette procédure, l'A.R.H. se substitue alors à l'hôpital pour demander sa mise en œuvre. Le directeur de l'A.N.A.E.S. fournit aux directeurs des A.R.H. toutes les informations quantitatives et qualitatives sur les programmes d'accréditation en cours dans les établissements de santé de leur région.

La présentation du mode de financement du service public hospitalier, en particulier les ordonnances de 1996 et le P.M.S.I., nous a semblé une étape nécessaire avant d'appréhender, d'un point de vue économique, le comportement des acteurs de ce service public. Comment l'hôpital peut-il maintenir sa D.G.F. ou encore justifier une sur-dotation ? Avant de proposer un modèle d'agence dans lequel interviendraient les hôpitaux, l'A.R.H. et le Ministère de la Santé, nous présentons dans la section suivante une revue de la littérature de la théorie de l'agence.

²⁹ Loi hospitalière du 31 juillet 1991.

SECTION 2 : DÉFINITIONS ET PRE-REQUIS DE LA THÉORIE DE L'AGENCE

Depuis ses premières définitions au début des années 1970 jusqu'aux modélisations actuelles, le concept de la relation d'agence ou de mandat s'est considérablement enrichi. Nous définissons la relation d'agence en information complète, parfaite et certaine, puis en asymétrie d'information. Nous effectuons alors une revue de la littérature des modèles de la relation d'agence.

1. L'information et la relation d'agence

1.1. Les premières définitions de la relation d'agence

Sans étudier explicitement la relation d'agence, la littérature économique des années 1960 et du début des années 1970 relative à l'aléa moral [Arrow, 1970] concerne déjà les problèmes d'agence. C'est S. Ross, qui fournit la première définition de la théorie de l'agence en économie [Ross, 1973]. Il s'agit d'une relation entre une partie informée (l'agent) et une partie non informée (le principal), l'objectif du principal étant de proposer à l'agent un contrat ayant deux propriétés. D'une part le contrat doit être acceptable pour l'agent ; d'autre part il doit fournir au principal une utilité maximale. Un optimum de premier rang est atteint si le principal et l'agent ont les mêmes préférences ou si leurs préférences divergent mais le principal contrôle parfaitement l'action de l'agent. Le principal peut ainsi ordonner à l'agent de choisir l'action la plus efficace. Toutefois, un conflit entre le principal et l'agent qui cherchent chacun à maximiser leur utilité espérée, ou une mauvaise observation de l'action de l'agent par le principal, conduit à un optimum de second rang. Ce sont donc, en général, ces derniers qui sont obtenus. La difficulté de contrôler l'action ou le comportement d'un individu est examinée en détail par M. Spence et R. Zeckhauser dans le domaine de l'assurance, préalablement aux travaux de S. Ross sur la théorie de l'agence [Spence et *al.*, 1971]. Pour S. Ross, la théorie de l'agence présente d'autant plus d'intérêt, que les décisions sont prises en univers incertain, c'est-à-dire que l'état de la Nature est une variable aléatoire notée θ . La rémunération de l'action a de l'agent par le principal, notée w , s'écrit alors $w(a, \theta)$. Le fait, que l'agent (ou le principal) puisse détenir une information différente du principal (ou de l'agent) sur l'état de la Nature, constitue pour S. Ross la raison d'être de la relation d'agence.

Quelques années plus tard, J.C. Jensen et W.K. Meckling appliquent la théorie de l'agence à la théorie de la firme [Jensen et *al.*, 1976]. Ils constatent un foisonnement de la littérature économique sur la théorie de la firme. Toutefois, selon J.C. Jensen et W.K. Meckling, cette littérature concerne davantage les théories des marchés sur lesquels les firmes exercent un rôle important. La firme reste trop souvent représentée comme une « boîte noire » [Marshall, 1920] ou comme une technologie transformant les intrants en extrants sur un marché concurrentiel [Hicks, 1981]. J.C. Jensen et W.K. Meckling notent, toutefois, une volonté de construire une théorie de la firme basée sur d'autres modèles que la maximisation du profit [Simon, 1955 et 1959 ; Penrose 1958 ; Baumol, 1959 ; Williamson, 1964 ; Williamson, 1970]. Pour O.E. Williamson notamment, les modèles des théories de la firme basés sur la maximisation du profit sont inadaptés pour expliquer le comportement managerial des grandes firmes [Williamson, 1975]. C'est dans ce contexte, bien décrit par certaines revues de la littérature comme celles de A.A. Alchian [Alchian, 1965], de F. Machlup [Machlup, 1967] ou encore de R.M. Cyert et C.L. Hedrick [Cyert et *al.*, 1972], que J.C. Jensen et W.K. Meckling appliquent la théorie de l'agence au sein de la théorie de la firme, en particulier les relations entre les actionnaires et les dirigeants. Ils assimilent la relation d'agence à un contrat par lequel l'actionnaire (principal ou mandant) confie au dirigeant (agent ou mandataire) la gestion de l'entreprise. Si chacune des parties maximise son utilité, il existe alors d'excellentes raisons de penser, que l'agent n'agira pas dans le meilleur intérêt du principal. L'actionnaire souhaite, par exemple maximiser les profits et les redistribuer le plus possible sous forme de dividendes, tandis que le dirigeant pourrait préférer autofinancer les investissements voire maximiser sa rémunération. La structure du capital de l'entreprise est alors définie par rapport aux décisions prises par le dirigeant ; or ces dernières n'optimisent pas nécessairement les fonctions d'utilité des actionnaires. Le principal doit alors maîtriser l'action de l'agent par de l'incitation et des contrôles. J.C. Jensen et W.K. Meckling mettent l'accent sur les coûts inhérents à cette maîtrise désignés sous le terme de coûts d'agence [Jensen et *al.*, 1976]. Ils analysent les coûts d'agence générés par les accords contractuels entre les propriétaires et les dirigeants de la firme. Leur niveau dépendrait du cadre réglementaire et de l'ingéniosité humaine à inventer des contrats optimaux. La nature des risques reste propre à chaque partie, en particulier la perte des apports financiers pour l'actionnaire et le licenciement pour le dirigeant. Le contrat opère un partage optimal du risque entre le principal généralement neutre vis à vis du risque et l'agent le plus souvent averse au risque. Un agent est neutre au risque au sens de Wakker, si quelle que soit sa

situation très risquée ou au contraire peu risquée, un point de probabilité supplémentaire donne une variation identique de son utilité.³⁰ Avec l'introduction de la théorie de l'agence au sein de la théorie de la firme, c'est la structure interne de la firme qui est étudiée avec ses problèmes de contrôle, d'incitation, de coordination, de communication, etc... L'évolution des concepts d'information, d'incertitude et de rationalité des individus permettent une complexité croissante des modèles d'agence.

1.2. Une relation d'agence basique en information parfaite, complète et certaine

E. Rasmusen développe un jeu dans le cadre du marché du travail entre un employeur (le principal) et un employé (l'agent) [Rasmusen, 1995 ; p. 169]. Selon cet auteur, l'information est parfaite si tous les ensembles d'information ne contiennent qu'un nœud unique. Autrement dit, aucun joueur n'aura de doute sur ce qui s'est passé dans le jeu jusque là. L'information est complète si la Nature ne joue pas en premier ou si son mouvement initial est observé par tous les joueurs. Enfin, l'information est certaine si la Nature ne joue pas après que tous les joueurs aient joué [Rasmusen, 1995 ; p. 45]³¹. Soient $q(e)$ la production en valeur fonction croissante du niveau d'effort e de l'agent, e^* le niveau d'effort optimal de l'agent, w le salaire offert à l'agent, \tilde{w} le salaire offert à l'agent qui satisfait la contrainte de participation et \underline{U} l'utilité de réservation de l'agent c'est-à-dire l'utilité de l'agent en deçà de laquelle il refuse le contrat proposé par le principal.

La fonction d'utilité du principal, fonction croissante de la différence entre la production en valeur q et le salaire w s'écrit

$$V(q - w) \text{ avec } V' > 0 \text{ et } V'' < 0 \quad (1)$$

La fonction d'utilité de l'agent, fonction décroissante de l'effort e et croissante du salaire w s'écrit

$$U(e, w) \text{ avec } U' > 0 \text{ et } U'' < 0 \quad (2)$$

Bien souvent dans les modèles d'agence de ce type, les principaux et les agents ont des concurrents. En période de pénurie de main d'œuvre, la concurrence concerne plus largement les principaux et en période de chômage les agents. Si l'agent rejette le contrat alors son résultat est égal à son utilité de réservation \underline{U} et le résultat du principal est nul. Cette utilité

³⁰ Se référer au paragraphe 1.6. « Les comportements face au risque », p. 50.

³¹ Le lecteur trouvera plus de précisions sur la Nature de l'information dans le paragraphe 1.3. « L'information imparfaite, incomplète et incertaine », p. 46.

de réservation existe, même si le principal se trouve en situation de monopole car l'agent peut rester au chômage si le salaire proposé est trop faible $w < \underline{U}$; l'assurance chômage pouvant être un mécanisme d'incitation à préférer une situation de chômage [Rueff, 1925 ; Rueff, 1931]. En revanche, si l'agent accepte le contrat alors son résultat correspond à $U(e, w)$; le résultat du principal étant égal à $V(q - w)$.

L'ordre du jeu séquentiel est le suivant : le principal propose à l'agent un salaire w (phase 1), puis, l'agent accepte ou rejette le contrat (phase 2). Si l'agent accepte le contrat, il détermine alors son niveau d'effort e . Enfin, le résultat en valeur du principal est fonction de $q(e)$ avec $q' > 0$. Trouver l'équilibre revient donc à trouver le contrat, qui fournit au principal une utilité maximale tout en étant acceptable pour l'agent. Le principal doit aussi anticiper la réaction de l'agent aux incitations du contrat. L'agent doit recevoir un salaire \tilde{w} pour exercer un effort e , où $\tilde{w}(e)$ est défini comme le salaire qui satisfait la contrainte de participation

$$U(e, \tilde{w}(e)) = \underline{U} \quad (3)$$

Le problème du principal est alors de maximiser l'effort de l'agent au sein de sa fonction d'utilité

$$\text{Max}_e [V(q(e) - \tilde{w}(e))] \quad (4)$$

Les conditions de premier ordre sont

$$V'(q(e) - \tilde{w}(e)) \cdot \left(\frac{dq}{de} - \frac{d\tilde{w}}{de} \right) = 0 \quad (5)$$

ce qui implique que

$$\frac{dq}{de} = \frac{d\tilde{w}}{de} \quad (6)$$

D'après la contrainte de participation et le théorème de la fonction implicite on a

$$\frac{d\tilde{w}}{de} = - \left(\frac{\partial U / \partial e}{\partial U / \partial \tilde{w}} \right) \quad (7)$$

En combinant (5) et (6) on a

$$\frac{\partial U}{\partial \tilde{w}} \cdot \frac{dq}{de} = - \frac{\partial U}{\partial e} \quad (8)$$

Ainsi, au niveau d'effort optimal e^* , l'utilité marginale de l'agent, s'il conservait intégralement la valeur de la production marginale, serait égale à la désutilité marginale de l'effort.

Si nous supposons une parfaite compétition entre les employeurs, l'utilité de réservation \underline{U} est choisie telle que

$$U(e^*, q(e^*)) = \underline{U} \quad (9)$$

Le principal doit élaborer un contrat tel que l'agent choisisse le niveau d'effort optimal. Trois types de contrats, équivalents en information complète, s'offrent au principal :

Le contrat « forcé », où $w = w^*$ lorsque $e = e^*$ et $w = 0$ lorsque $e \neq e^*$ est une forte incitation pour contraindre l'agent à choisir $e = e^*$.

Le contrat « indexé », où $w = w^*$ lorsque $e \geq e^*$ et $w = 0$ lorsque $e < e^*$. L'agent bénéficie d'un salaire fixe lorsque le niveau d'effort est inférieur au niveau d'effort optimal (le salaire est même nul dans ce contrat) et d'un salaire majoré lorsque l'effort est supérieur à l'effort optimal. Si l'agent est inverse à l'effort, il choisit le niveau d'effort $e = e^*$.

Le contrat « linéaire », où $w(e) = \alpha + \beta e$ où α et β sont choisis tels que $w^* = \alpha + \beta e^*$ et la droite des contrats est tangente à la courbe d'indifférence $U = \underline{U}$.

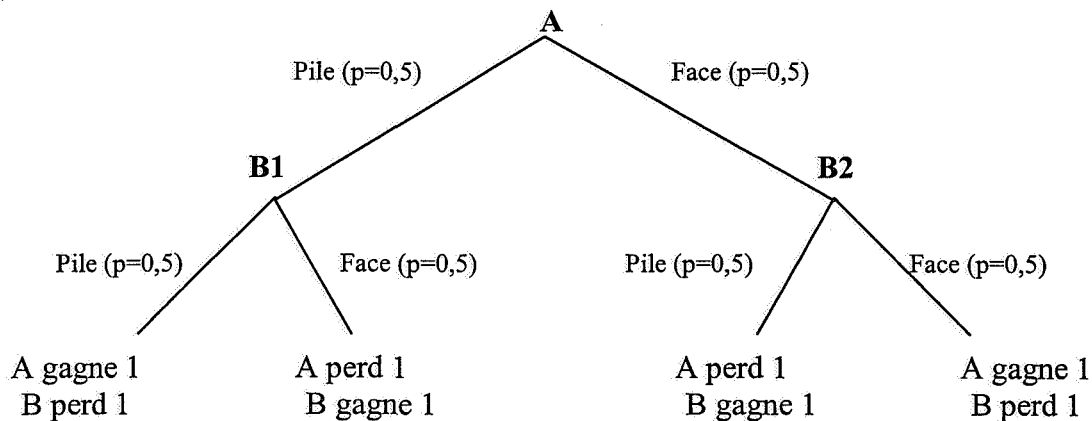
1.3. L'information imparfaite, incomplète et incertaine

Dans le cadre de jeux répétés, l'information parfaite versus imparfaite se distingue fondamentalement de l'information complète versus incomplète [Harsanyi, 1995]. D. Fudenberg et J. Tirole définissent l'information parfaite lorsque tous les ensembles d'information ne contiennent qu'un nœud unique. « *All the information sets are singletons. In a game of perfect information, players move at a time, and each player knows all previous moves when making his decision* » [Fudenberg et al., 1995 ; p.80]. En information imparfaite, l'asymétrie d'information porte sur le passé. Au moins un des individus (ou joueurs) ne sait pas exactement ce qui s'est déroulé antérieurement.

En information incomplète, l'incertitude porte sur l'identité du joueur. Cette identité comprend la fonction d'utilité et les moyens d'action ou les stratégies du joueur (le plan d'action qui spécifie son comportement au cours du jeu). « *When some players do not know the payoffs of the others, the game is said to have incomplete information* » [Fudenberg et al., 1995 ; p.80]. Ainsi, la « couinche » est un jeu à information parfaite (les levées précédentes sont connues des 4 joueurs) mais incomplète (les mains des 3 autres joueurs demeurent inconnues, sauf en cas de tricherie !). Prenons un autre jeu dans lequel un joueur A puis un

joueur B jettent successivement une pièce avec une probabilité $p=0,5$ d'obtenir pile et $p=0,5$ d'obtenir face. Si le résultat est pile-pile ou face-face A gagne 1 et B perd 1. Si le résultat est pile et face ou inversement, A perd 1 et B gagne 1. Ce jeu peut être représenté à partir de la figure 6. En information parfaite, B connaît le résultat de A et sait s'il se trouve en B1 ou B2 (l'ensemble d'information est un singleton). En information imparfaite, B ne connaissant pas le résultat de A, ne sait pas s'il se trouve en B1 ou B2 (l'ensemble d'information est une paire). En information complète, A et B connaissent toutes les distributions de probabilité et les résultats possibles du jeu. En information incomplète, au moins une distribution de probabilité ou un résultat possible reste inconnu de l'un des joueurs. Toutes les circonstances dans lesquelles le jeu est joué ne sont pas décrites pour tous les joueurs.

Figure 6
Représentation du jeu sous forme d'arbre



E. Rasmusen propose et définit, quant à lui, quatre types d'informations présentes en théorie des jeux :

(1) Concernant l'information parfaite versus imparfaite, il rejoint l'analyse de D. Fudenberg et J. Tirole selon laquelle chaque ensemble d'information ne contient qu'un nœud unique. « *In a game of perfect information each information set is a singleton. Otherwise the game is one of imperfect information* » [Rasmusen, 1995 ; p.45].

(2) En revanche, E. Rasmusen différencie l'information complète de l'information incomplète sur la base d'une observation commune ou non des modifications des états de la Nature « *In a game of incomplete information, Nature moves first and is unobserved by at least one of the players. Otherwise the game is one of complete information* » [Rasmusen, 1995 ; p.47].

(3) Puis, il complète sa typologie en distinguant l'information symétrique versus asymétrique : « *In a game of symmetric information, a player's information set at any node where he chooses an action, or an end node contains at least the same elements as the information sets of every other player. Otherwise the game is one of asymmetric information* » [Rasmusen, 1995 ; p.47].

(4) Enfin, le quatrième type d'information concerne l'information certaine versus incertaine : « *A game of certainty has no moves by Nature after any player has moved. Otherwise the game is one of uncertainty* » [Rasmusen, 1995 ; p.45].

Dans le cadre de la relation d'agence, un des joueurs (l'agent) peut détenir soit une information sur un état de la Nature, soit une information passée, soit un élément de sa propre identité.

1.4. Les degrés d'incertitude

Si l'asymétrie d'information repose sur le fait qu'un individu ne connaît pas les caractéristiques de l'autre, l'incertitude est quant à elle liée à la probabilité que l'état de la Nature se modifie au cours d'une période. L'état de la Nature concerne par exemple l'environnement de la firme, c'est-à-dire la structure du marché, l'état de la demande, l'importance de la concurrence, etc... Il convient donc de distinguer l'incertitude de l'asymétrie d'information. L'incertitude se décline en cinq degrés, de l'ignorance totale à la connaissance commune. L'individu en situation d'ignorance totale (ou radicale) ignore sa propre ignorance. Il est incapable de se représenter l'ensemble des états de la Nature, puisqu'il ne peut même pas les concevoir. Le second degré concerne l'individu en situation d'ignorance. Conscient de son ignorance, il possède la faculté d'identifier l'ensemble des états de la Nature. L'incertitude probabilisable avec des probabilités subjectives constitue le troisième degré. L'individu connaît l'ensemble des états de la Nature et détient des croyances sur celui-ci. L'incertitude probabilisable avec des probabilités objectives forme le quatrième degré. L'individu possède une capacité d'apprentissage et révisé ses croyances. Enfin, la connaissance commune repose sur une incertitude probabilisable avec des probabilités objectives communes aux individus, qui se représentent donc l'incertain de façon identique.

1.5. La rationalité des agents économiques

Les capacités cognitives, la Nature et le degré d'incertitude, le mode de prise de décision et les objectifs poursuivis par les agents économiques sont au cœur du principe de

rationalité. Dans le cadre de la rationalité substantielle, hypothèse fondamentale du modèle néoclassique, les individus disposent d'une information parfaite et complète. La rationalité substantielle décrit le comportement optimisateur des individus, qui est à l'origine de leur prise de décision. Ainsi, en anticipant les conséquences sur leur intérêt personnel de l'ensemble des choix qui s'offrent à eux, ils maximisent leur satisfaction. Cette hypothèse suppose donc, que l'individu est doté de capacités de calcul, qui lui permettent d'anticiper les conséquences de ses décisions. Elle suppose aussi, que l'individu (entreprise ou ménage) adopte un comportement individualiste et que son seul objectif soit l'obtention d'un gain en termes d'utilité, de revenu ou de profit. Le programme de l'agent économique consiste alors à maximiser une fonction objectif sous des contraintes exogènes, imposées par le monde extérieur. La rationalité substantielle peut être contrainte ou limitée, lorsque l'information devient imparfaite et/ou incomplète ou que le degré d'incertitude augmente. Dans ce contexte, les anticipations des agents nécessitent des capacités de calculs beaucoup plus importantes soit pour prendre en compte la stratégie des autres acteurs, soit pour intégrer les coûts de recherche de l'information. Comme le souligne K.J. Arrow, l'abandon de l'hypothèse d'information parfaite conduit à faire peser une menace sur le concept même de rationalité, parce que la perception des autres, en particulier leur rationalité, devient un élément de notre propre rationalité. [Arrow, 1985a]. Dans cette configuration, l'enjeu du célèbre concours de beauté de J.M. Keynes nous apparaît pleinement [Keynes, 1971 (édition française)]. Dans le cadre de la rationalité limitée, les agents économiques ne maximisent pas des fonctions traduisant leurs préférences mais ils se fixent des buts qu'ils essaient au mieux de satisfaire [Simon, 1964]. Compte tenu des limites de la rationalité substantielle en termes de capacités cognitives des individus et de représentation de la réalité [Giordano, 1991], H. Simon propose le concept de rationalité procédurale [Simon, 1982]. Il prend ainsi le contrepied du paradigme de la rationalité substantielle en montrant qu'il néglige tout ce qui concerne la psychologie du décideur. La rationalité procédurale consiste à sélectionner une décision parmi un éventail non nécessairement exhaustif d'alternatives, en fonction d'un système de valeur et d'information qui permet d'évaluer les conséquences de cette décision [Favereau, 1989]. Même si l'individu disposait de toute l'information, ses capacités à traiter cette information abondante seraient insuffisantes. Le concept de rationalité procédurale est importé au sein de la relation d'agence [Tirole, 1986].

1.6. Les comportements face au risque

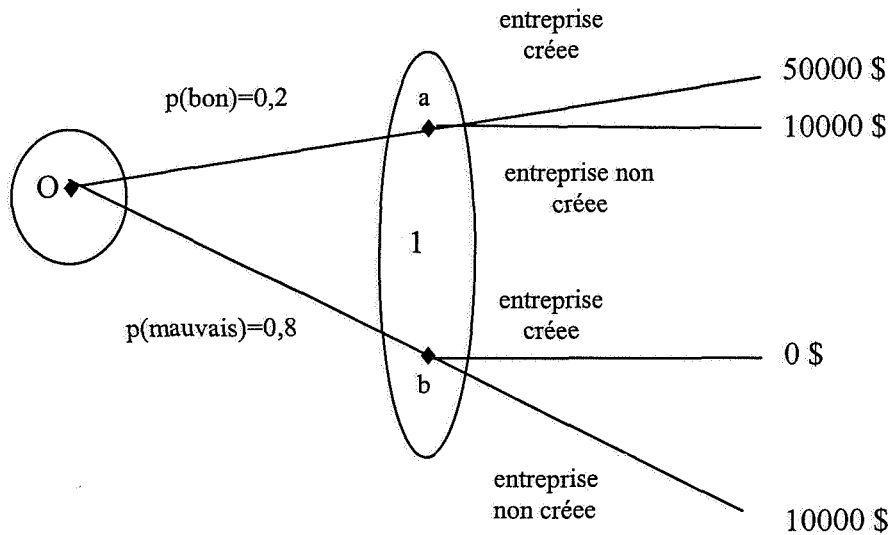
1.6.1. L'approche de K. J. Arrow et J. W. Pratt

Comme le fait remarquer K J Arrow, il existe déjà une multitude d'approches de la notion de risque au début des années 1950 [Arrow, 1951]. Toutefois, l'introduction du coefficient d'aversion au risque par K.J. Arrow et J.W. Pratt, dans le cadre d'une démarche numérique, marque une importante avancée [Arrow, 1963 ; Pratt, 1964]. Bien d'autres ont, par la suite, largement contribué à développer le concept de risque, en particulier G. Hanoch et H. Levy, M. Rothschild et J.E. Stiglitz pour les comportements plus risqués mais aussi S Ross, M.J. Machina en ce qui concerne l'aversion au risque [Hanoch et *al.*, 1969 ; Rothschild et *al.*, 1970 ; Ross, 1981 ; Machina, 1982].

Un joueur peut être averse au risque (risquophobe), neutre au risque ou encore preneur de risque (risquophile). Si un joueur est neutre au risque, alors un dollar certain est le même qu'un dollar espéré, autrement dit avoir la certitude d'obtenir 1000 \$ a exactement la même valeur qu'une probabilité de distribution ayant une valeur espérée de 1000 \$. Le risque n'affecte pas la décision d'un joueur neutre au risque, tant qu'il n'affecte pas la valeur espérée. Si le joueur est averse au risque, alors, obtenir un dollar avec certitude est préférable à un dollar espéré. Ce joueur évite le risque à moins que la chance soit suffisamment en sa faveur. Inversement, un joueur risquophile préfère un dollar espéré à un dollar sûr. Ce joueur prend des risques à moins que la chance soit suffisamment contre lui. Les fonctions d'utilité sont spécifiques à la position du joueur face au risque, concaves pour les risquophobes, linéaires pour les joueurs neutres au risque, convexes pour les risquophiles. Prenons l'exemple du jeu « small business » à information imparfaite et comprenant un joueur [Gardner, 1995]. En information imparfaite, un joueur doit décider de créer ou de ne pas créer une entreprise, sachant qu'une création d'entreprise comporte toujours un risque. Soit $p(\text{bon})$, la probabilité d'avoir une conjoncture économique favorable et donc une entreprise qui prospère et $p(\text{mauvais})$, la probabilité d'avoir une mauvaise conjoncture économique et de déposer le bilan. Le joueur ne connaît pas l'état de la conjoncture lorsqu'il prend la décision. Les données du jeu sont les suivantes : le joueur dispose de 10000 \$ pour créer son entreprise. S'il ne la crée pas, il conserve ses 10000 \$. S'il la crée et que la conjoncture économique est bonne, il réalisera alors un profit de 50000 \$. Enfin, s'il la crée et que la conjoncture est défavorable, alors son résultat sera nul. Les trois résultats se représentent facilement sous la

forme d'un arbre (figure 7).

Figure 7
Arbre du jeu « small business »



Source : [Gardner, 1995, p. 18]

Nous sommes bien en information imparfaite car en 1 l'ensemble d'information est une paire $\{a,b\}$, puisque le joueur ne sait pas où s'il se situe. Pour résoudre ce jeu, le joueur doit se demander si conserver ses 10000 \$ est préférable à la distribution de probabilité suivante : 50000 \$ avec une probabilité de 0,2 et 0 \$ avec une probabilité de 0,8.

$$u_1(10000) = 10000\$$$

$$\begin{aligned} Eu_1(\text{entreprise créée}) &= p(\text{bon}) \times u_1(50000) + p(\text{mauvais}) \times u_1(0) \\ &= 0,2 \times 50000 + 0,8 \times 0 = 10000\$ \end{aligned}$$

(1) Pour le joueur neutre au risque, conserver ses 10000 \$ avec certitude ou créer l'entreprise dont l'espérance de la distribution de probabilité est également de 10000 \$ a exactement la même valeur. Ces deux choix se situent sur la même courbe d'indifférence. Seul le joueur neutre au risque exprime directement ses dollars en utilité, puisque sa fonction d'utilité est linéaire.

(2) Le joueur averse au risque doit avant tout transformer de façon non linéaire ses dollars en utilité. Prenons la fonction suivante :

$$u_1(m_1) = \frac{m_1^a}{a} \quad \text{si } a \neq 0$$

$$u_1(m_1) = \log(m_1) \quad \text{si } a = 0$$

Si $a < 1$, le joueur est averse au risque et plus a est grand et moins le joueur est averse au risque. Le joueur averse au risque ne crée pas son entreprise. En effet, pour $a = 0,5$ par exemple, nous obtenons la fonction d'utilité concave suivante :

$$u_1(m_1) = \frac{m_1^{0,5}}{0,5} = 2(m_1^{0,5})$$

Nous transformons maintenant les résultats exprimés sous forme monétaire (dollars) en utilité avec, rappelons-le, $a = 0,5$.

$$u_1(10000) = 2(10000)^{0,5} = 200$$

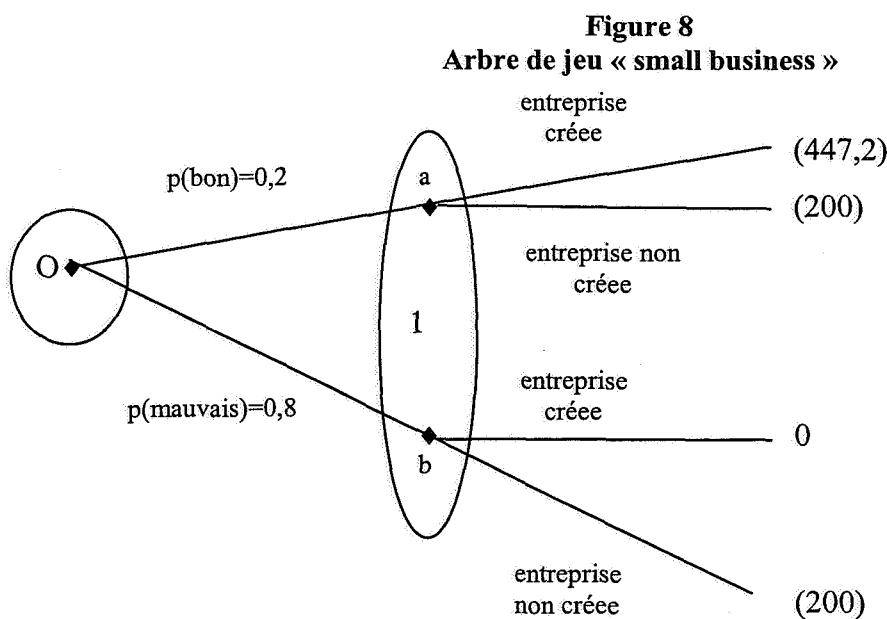
$$u_1(50000) = 2(50000)^{0,5} = 447,2$$

Les utilités espérées deviennent :

$$Eu_1(\text{créer l'entreprise}) = 0,2 \times 447,2 + 0,8 \times 0 = 89,4$$

$$Eu_1(\text{ne pas créer l'entreprise}) = 2(10000)^{0,5} = 200$$

Ce joueur, adverse au risque, choisit de ne pas créer l'entreprise, puisque l'utilité espérée « ne pas créer l'entreprise » est supérieure à l'utilité espérée « créer l'entreprise » (200 contre 89,4). Les résultats sont présentés dans la figure 8.



Source : [Gardner, 1995, p. 21]

(3) Le joueur preneur de risque ou risquophile doit également transformer de façon non linéaire ses dollars en utilité avec $a > 1$, puisque sa fonction d'utilité est convexe. Si nous reprenons la fonction précédente, pour $a = 2$ elle s'écrit :

$$u_1(m_1) = \frac{m_1^2}{2}$$

$$u_1(10000) = \frac{10000^2}{2} = 5 \times 10^7$$

$$u_1(50000) = \frac{50000^2}{2} = 1,25 \times 10^9$$

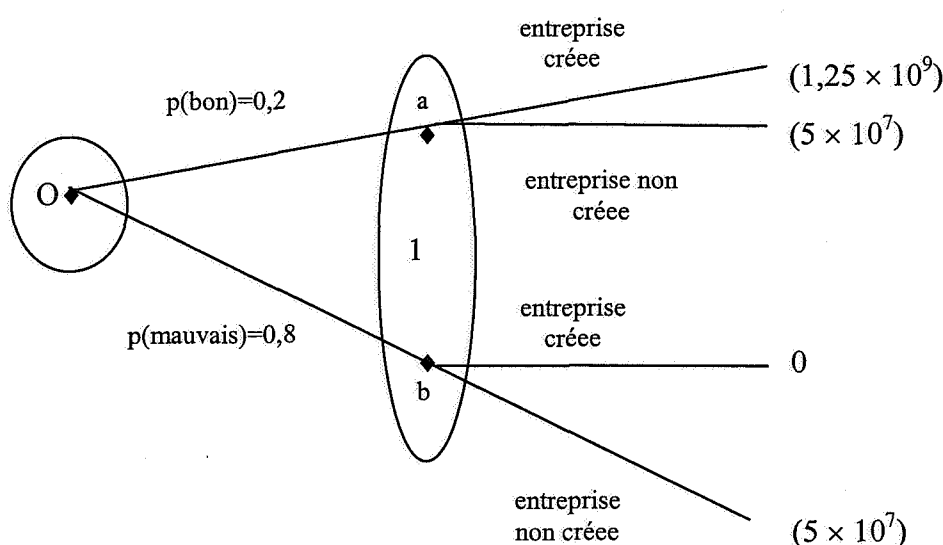
Les utilités espérées deviennent :

$$Eu_1(\text{créer l'entreprise}) = 0,2 \times (1,25 \times 10^9) + 0,8 \times 0 = 250 \times 10^6$$

$$Eu_1(\text{ne pas créer l'entreprise}) = \frac{10000^2}{2} = 50 \times 10^6$$

Ce joueur, risquophile choisit de créer l'entreprise, puisque l'utilité espérée « créer l'entreprise » est supérieure à l'utilité espérée « ne pas créer l'entreprise » (250 millions contre 50 millions).

Figure 9
Arbre de jeu « small business »



Source : [Gardner, 1995, p. 22]

La notion de risque en probabilité constitue également un apport théorique important.

1.6.2. L'approche de P. Wakker

La notion de risque en probabilité a été développée par P. Wakker [Wakker, 1994] et est clairement exposée dans l'ouvrage de O. Jokung-Nguéna [Jokung-Nguéna, 1998]. Néanmoins, nous nous appuyons dans ce paragraphe sur un document de travail élaboré par G. Duru [Duru, 2000].

Soit L , un ensemble dont les éléments sont appelés des lots. On appelle « loterie discrète finie » toute variable aléatoire discrète et finie dont le support de distribution de probabilité est un ensemble fini contenu dans L .

Soit X un ensemble d'objets sur lequel un agent économique k exerce des préférences décrites par une relation binaire \geq_k sur X définie par :

Quels que soient x_1 et x_2 appartenant à X , x_1 est préféré ou indifférent à x_2 pour l'agent k si et seulement si $x_1 \geq_k x_2$. On appelle \geq_k une « relation de préférence de l'agent k ». On suppose que X^2 est une structure neumanienne [Von Neumann et al. 1953], définie par une suite de fonctions $(h_n)_{n \in \mathbb{N}}$, qui sont des applications de $X^n \Delta(n)$ dans X avec :

$$\Delta(n) = \left\{ \alpha \in \mathbb{R}^n / \alpha = (\alpha_1 \dots \alpha_n), \alpha_i \geq 0, \sum \alpha_i = 1 \right\}$$

Nous définissons la fonction (h_2) sur $X^2 \Delta(2)$:

Soient x_1 et x_2 deux loteries discrètes finies caractérisées par leurs distributions de probabilité respectives $P_1 = (P_{11}, \dots, P_{1n})$ sur $\{I_{11}, \dots, I_{1n}\} \subset L$ et $P_2 = (P_{21}, \dots, P_{2m})$ sur $\{I_{21}, \dots, I_{2m}\} \subset L$. I_{11}, \dots, I_{1n} correspond aux lots de la loterie x_1 et I_{21}, \dots, I_{2m} aux lots de la loterie x_2 . On définit $h_2[x_1, x_2, (\alpha_1, \alpha_2)]$ comme la loterie équivalente au jeu qui consiste à tirer au sort l'une des loteries x_1, x_2 selon les probabilités α_1, α_2 et jouer alors à la loterie qui est gagnée.

(1) Un agent économique k est neutre au risque en probabilité si et seulement si quelles que soient les loteries x_1 et x_2 et quel que soit α appartenant à $\Delta(2)$

$$x_1 \geq_k x_2 \Rightarrow x_1 \geq_k h_2(x, y, \alpha) \geq_k x_2$$

(2) L'agent économique est averse au risque en probabilité si et seulement si quelles que soient les loteries x_1 et x_2 et quel que soit α appartenant à $\Delta(2)$

$$x_1 \geq_k x_2 \Rightarrow x_1 \geq_k h_2(x, y, \alpha)$$

(3) L'agent économique prône le risque en probabilité si et seulement si quelles que soient les loteries x_1 et x_2 et quel que soit α appartenant à $\Delta(2)$

$$x_1 \geq_k x_2 \Rightarrow h_2(x, y, \alpha) \geq_k x_2$$

Nous introduisons maintenant l'asymétrie d'information au sein de la relation d'agence, c'est-à-dire les notions de risque moral et de sélection adverse entre autres.

2. L'asymétrie d'information au sein de la relation d'agence

La théorie de l'agence s'inscrit avant tout dans un cadre néo classique. Toutefois, l'asymétrie d'information implique la remise en cause de cet environnement walrasien basé sur trois hypothèses majeures, la rationalité substantielle, une économie d'allocation de ressources et une coordination marchande des décisions individuelles [Walras, 1874].

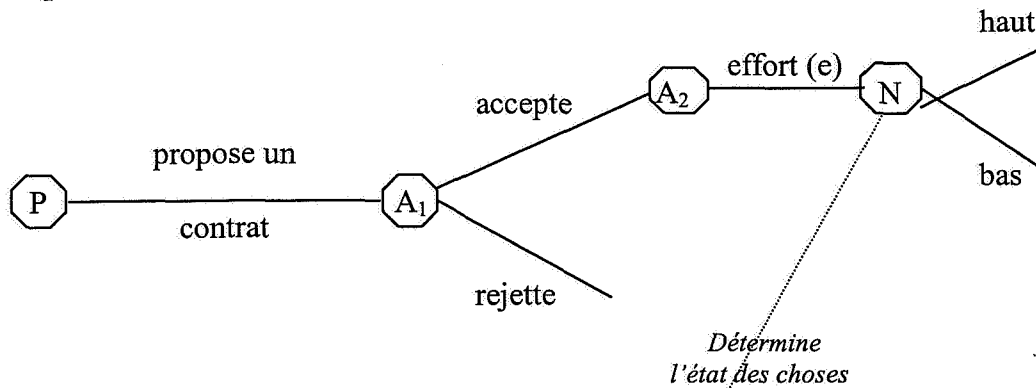
2.1. L'introduction du risque moral au sein de la relation d'agence

Comme nous l'avons vu, la théorie de l'agence étudie les contrats, qui permettent au mandataire (le principal) d'inciter le mandant (l'agent) à agir dans l'intérêt du mandataire. Mais l'action de l'agent n'est pas toujours observable par le principal ou à un coût prohibitif. On parle dans ce cas de risque moral ou d'aléa moral [Macho-Stadler et al., 1997]. L'asymétrie d'information est alors post contractuelle et concerne l'effort fourni par l'agent, ce dernier pouvant être incité à adopter un comportement déviant par rapport aux termes du contrat préalablement accepté par les deux parties. Prenons l'exemple d'une firme, qui peut assurer l'intégralité de son processus de production ou sous-traiter une partie de celui-ci pour réduire ses coûts de production. Si la firme choisit la relation d'agence en sous-traitant une partie de son processus de production, elle s'expose au risque moral, puisqu'elle perd le contrôle de ce processus [Lewis et al., 1990]. Un fois le contrat signé, la firme contrôle moins la qualité et les délais de production des biens intermédiaires. L'imputation des responsabilités quand à l'effort fourni impose de distinguer le risque moral avec une *action cachée* du risque moral avec une *information cachée*. Dans le cas du risque moral avec une

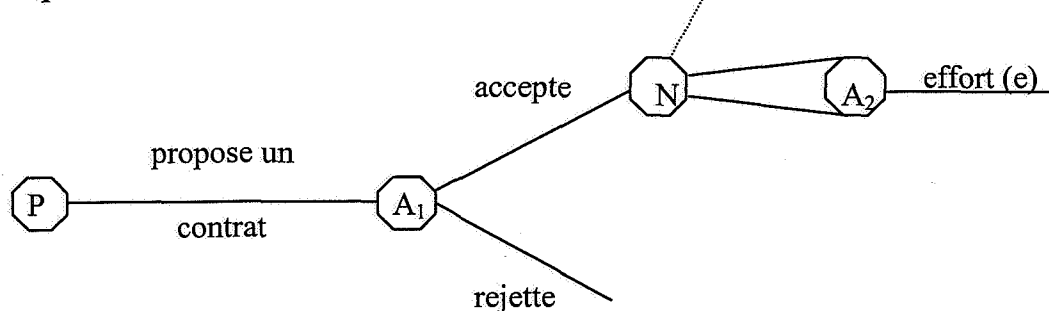
action cachée, l'agent accepte un contrat proposé par le principal, puis détermine son niveau d'effort avant de connaître l'état de la Nature (les variables non maîtrisées comme la conjoncture économique). Autrement dit, après l'établissement de la relation contractuelle, le comportement de l'agent n'est pas une variable vérifiable de la relation, celui-ci ne peut donc être inclus dans les termes du contrat. En terme de jeu cela revient à dire que l'agent joue avant la Nature. Dans le cas du risque moral avec de l'information cachée, l'agent connaît l'état de la Nature lorsqu'il fixe son niveau d'effort. Cette information lui permet de déterminer son niveau d'effort adéquat, cette observation n'étant pas observable ou vérifiable par le principal. Dans ce cas, la Nature joue avant l'agent. Nous pouvons résumer cela à l'aide des arbres de jeu présentés dans la figure suivante.

Figure 10
Action et information cachées

Risque moral avec de l'action cachée



Risque moral avec de l'information cachée



Avec P : le principal ; A : l'agent ; N : la Nature.

L'aléa moral complexifie les modèles d'agence puisque des programmes de maximisation avec des conditions du second ordre sont nécessaires. En reprenant l'exemple de

E. Rasmusen sur la relation de travail, nous pourrions obtenir le programme de maximisation suivant :

En supposant la fonction d'utilité de l'agent séparable, on a

$$U(w, e) = U(w) - d(e) \quad (1)$$

$$\text{avec } u'(w) > 0 \quad u''(w) \leq 0 \quad \text{et} \quad d'(e) > 0 \quad d''(e) > 0$$

L'utilité de l'agent croît avec sa rémunération w et décroît avec son action ou son effort e . La désutilité marginale est croissante à l'effort tandis que l'utilité marginale de la rémunération est décroissante. Le contrat doit aussi être acceptable pour l'agent c'est-à-dire lui fournir une utilité supérieure à son utilité de réservation U . Le niveau de rationalité individuelle de l'agent étant exogène et normalisé à zéro, on a

$$U(w, e) \geq 0 \quad (2)$$

La fonction d'utilité du principal est

$$V(q - w) \quad \text{avec } V' > 0 \quad \text{et} \quad V'' < 0 \quad (3)$$

Comme l'effort e n'est pas observable par le principal, la rémunération w de l'agent est fonction uniquement de la production q . Par ailleurs, le principal sait que l'agent choisit son action dans son meilleur intérêt, c'est-à-dire

$$e \in \arg \max [u(w(q)) - d(e')] f(q, e) dx \quad (4)$$

Les formules (2) et (4) forment les contraintes et (4) est remplacée par la condition du premier ordre. Le contrat $(w(q), e)$ proposé par le principal à l'agent est alors solution du programme

$$\begin{aligned} & \max \int [q - w(q)] f(q, e) dx \\ & \text{avec } \int [u(w(q)) - d(e')] f(q, e) dx \geq 0 \quad \text{pgr (1)} \\ & \text{et } \int u(w(q)) f(w, e) dx - d'(e) = 0 \end{aligned}$$

L'aléa moral conduit à un optimum de second rang, puisque compte tenu de l'asymétrie de l'information, l'observation de l'action de l'agent par le principal est partielle. E. Rasmusen distingue plusieurs types de contrats [Rasmusen, 1995] :

(1) Le principal propose à l'agent un salaire fixe en information incomplète mais certaine : Ce jeu convient lorsque le principal, en asymétrie d'information, ne peut déterminer le salaire de l'agent ni à partir du niveau d'effort, ni à partir de la production en valeur. C'est le cas, par exemple, de certains services comme l'enseignement. Si aucun élément ne permet de déterminer le salaire, la relation d'agence ne peut être résolue. En général, dans les jeux répétés, d'autres variables comme la réputation entrent en compte.

(2) Le principal propose à l'agent un salaire variable en information incomplète mais certaine : Le principal n'observe pas le niveau d'effort de l'agent, mais la production réalisée. Il propose à l'agent non pas un salaire w mais une fonction $w(q)$. Malgré l'inobservation de l'effort, le principal peut déterminer un niveau d'effort optimal e^* , qui permet de fournir une production optimale en valeur q^* avec $q = q(e^*)$. Le contrat doit alors inciter l'agent à choisir e^* . La non observation de l'effort n'est donc pas un problème au sein de la relation d'agence tant que le contrat peut être déterminé par un élément observable et parfaitement corrélé à l'effort. Or, en introduisant l'information incertaine, cette corrélation disparaît puisque la Nature peut jouer après que chacun des joueurs ait joué.

(3) Le principal propose à l'agent un salaire variable en information incomplète et incertaine : Le principal n'observe là encore pas l'effort mais la production en valeur et spécifie le contrat $w(q)$. La production est fonction de l'effort fourni par l'agent et de l'état de la Nature $q(e, \theta)$: avec $\theta \in \mathfrak{R}$. Les changements induits par l'état de la Nature interviennent lorsque l'agent a déterminé son niveau d'effort, l'agent ne revenant pas sur ce dernier car nous serions en présence d'aléa moral. Comme l'information est incertaine, l'effort n'apparaît plus clairement lors de l'observation de la production. Un niveau de production donné peut être alors atteint avec différents niveaux d'effort et, contrairement au jeu précédent, le principal ne peut plus déduire que $e = e^*$. Et, même si $e = e^*$, le principal n'est plus assuré d'obtenir q^* du fait de θ . L'information incertaine introduit un nécessaire compromis entre inciter l'agent à réaliser l'effort e^* et se prémunir du risque lié à l'état de la Nature θ .

R. Lal fournit une illustration intéressante de l'aléa moral dans le cadre d'une relation d'agence au sein du département marketing d'une firme [Lal, 1986]. Le directeur des ventes (le principal) peut imposer un prix de vente à l'agent commercial (l'agent). Il peut aussi considérer que l'agent commercial détient une meilleure information sur le marché et lui déléguer la fixation du prix de vente. A partir de la relation d'agence, R. Lal démontre que si le principal détient autant d'information sur le marché que l'agent commercial, alors il n'a aucune raison de déléguer la décision en matière de prix de vente à l'agent. Au contraire en situation d'asymétrie d'information, une telle délégation peut permettre à la firme de s'adapter plus rapidement par les prix, à l'évolution du marché. Ceci est d'autant plus vrai en cas, de forte élasticité-prix des biens produits, de pouvoir de négociation important des clients ou

encore de produits périssables. Mais le principal doit inciter l'agent, à partir d'un plan de compensation, à pratiquer des prix de vente plus élevés grâce à des efforts plus soutenus. R. Lal montre qu'une telle délégation accentue indubitablement le risque moral.

2.2. L'introduction de la sélection adverse au sein de la relation d'agence

On parle de sélection adverse ou anti sélection lorsque l'agent possède des informations privées au moment de contracter avec le principal [Macho-Stadler et al., 1997]. L'agent détient une telle information avant même de choisir son action, en l'occurrence, déterminer son niveau d'effort. Prenons l'exemple d'une firme, qui décide de sous-traiter une partie de son processus de production pour minimiser ses coûts. Cette décision entraîne une relation d'agence avec non seulement du risque moral mais aussi de la sélection adverse [Riordan et al., 1987]. Avant la signature du contrat, la firme donneur d'ordre (le principal) ne connaît pas toutes les caractéristiques du sous-traitant (l'agent), en particulier ses compétences, ses contrats avec d'autres firmes, etc... Le revenu du principal q n'est pas une variable aléatoire mais une fonction de l'action (ou de l'effort) e de l'agent et de l'état de la Nature θ . L'agent connaît l'état de la Nature avant de choisir son action. Le principal cherche alors à connaître θ , information privée de l'agent. Soit θ^0 , la valeur de θ que l'agent communique au principal. Le principal demande alors à l'agent de choisir l'action $e(\theta^0)$ et lui verse un salaire $w(\theta^0)$, sachant que l'agent peut mentir sur θ^0 .

Nous faisons l'hypothèse que θ prend les valeurs θ^1 ou θ^2 , avec $\theta^1 < \theta^2$. De plus, le principal est neutre au risque. La première étape consiste alors pour le principal à déterminer un contrat, qui incite l'agent à dire la vérité sur θ , en s'appuyant sur le principe de révélation de R. Myerson [Myerson, 1979 et 1982]. Soit w_1 , la rémunération versée à l'agent lorsque $\theta^0 = \theta^1$ et w_2 , la rémunération versée à l'agent lorsque $\theta^0 = \theta^2$. Supposons que lorsque θ^1 est la vraie valeur de θ , l'agent préfère transmettre au principal θ^2 et obtenir l'utilité U_1 avec $U_1 = U(w_1)$. Selon R. Myerson, il existe un contrat élaboré par le principal, qui oblige l'agent à dire la vérité, sans détériorer la situation du principal et de l'agent de la forme

$$\max_{w_j, e_j} E [q(w_j, e_j) - w_j] \quad (\text{contrainte 1})$$

$$\text{sous } U_1(w_j) - U_2(e_j) \geq 0$$

$$U_1(w_j) - U_2(e_j) \geq U_1(w_i) - U_2(e_i) \text{ quand } \theta = \theta_j \text{ avec } i, j = 1, 2 \text{ et } i \neq j \quad (\text{contrainte 2})$$

avec $w_j = w(\theta_j)$ et $e_j = e(\theta_j)$

Tandis que la première contrainte garantit à l'agent son utilité de réservation quelque soit θ_j , la seconde incite l'agent à dire la vérité car l'utilité associée à la vérité est supérieure à l'utilité associée au mensonge.

2.3. Les conséquences de l'asymétrie d'information

L'asymétrie d'information, qui génère le pouvoir, suscite les comportements opportunistes [Perroux, 1975]. L'opportunisme peut se définir comme la recherche d'un intérêt personnel, qui comporte la notion de tromperie [Williamson, 1994]. Quant au pouvoir, R. Dahl le définit comme la capacité à affecter le comportement d'autres personnes [Dahl, 1957]. En d'autres termes, le pouvoir de A sur B correspond à la capacité de A d'obtenir que dans sa négociation avec B, les termes de l'échange lui soient favorables. Pour J. Tirole et J-J. Laffont, la maîtrise, de la communication et des flux d'information, constitue la principale source de pouvoir [Tirole, 1986; Laffont, 1988]. Le pouvoir peut également provenir d'une compétence, d'un statut ou encore d'une position hiérarchique. L'asymétrie d'information peut décider le principal à contrôler l'agent ou à le faire contrôler par un tiers, dénommé superviseur. Au sein de cette nouvelle relation d'agence comprenant trois acteurs, naît le risque de collusion. Cette dernière peut être définie comme un accord conclu entre deux d'entre eux, sans que le troisième n'en soit informé.

2.4. Le signal et le dépistage pour atténuer l'asymétrie d'information

Pour atténuer l'asymétrie d'information, les parties disposent de signaux et de dépistages. Le signal émane de l'agent [Spence, 1973], tandis que le principal essaie de dépister les caractéristiques de l'agent. Ce dépistage implique des coûts d'agence, qui prennent la forme de coûts de surveillance, de coûts de motivation, voire de coûts d'opportunité. Ils varient selon le degré de confiance entre les différents acteurs. Choisir la relation d'agence lorsque cette dernière n'est pas imposée comme dans la relation médecin-patient suppose alors que le mandataire présente un avantage comparatif si important en termes de coûts, de qualité, etc... qu'il renverse les coûts induits par les problèmes d'asymétrie d'information associés à la relation d'agence.

Après avoir présenté la relation d'agence en information parfaite, complète et certaine, puis en

asymétrie d'information, nous effectuons une revue de la littérature des différents modèles d'agence.

3. Une typologie des modèles de la relation d'agence

Les travaux récents d'application à la santé de la théorie de l'agence se tournent vers des modèles plus complexes et mieux adaptés aux caractéristiques du système de santé dans lequel interviennent de multiples acteurs comme les représentants de l'autorité de tutelle, les producteurs de soins, le public ou encore les assureurs [Rochaix, 1997].

3.1. Les modèles à tâches multiples

Ces modèles comprennent un principal et un agent, le principal pouvant enjoindre à l'agent de réaliser simultanément plusieurs tâches. Le mandant fixe, par exemple, des objectifs à la fois en termes de coûts et de qualité [Holmström et *al.*, 1991]. Certains modèles recherchent le système de paiement optimal pour une organisation sanitaire donnée. Si l'objectif premier du producteur de soins est de maximiser son profit tandis que la tutelle privilégie d'abord la maîtrise des coûts puis la qualité, alors le système du paiement à la pathologie est optimal [Chalkley et *al.*, 1995]. En revanche, si le producteur de soins est plutôt altruiste alors un paiement, qui combine le remboursement partiel des coûts observés et le paiement forfaitaire *per capita* s'avère optimal [Ellis et *al.*, 1990]. Comme les incitations jouent en sens opposés sur ces deux variables essentielles (le coût et la qualité des soins), la régulation du système doit reposer sur des incitations et non sur des prix ou des contraintes. Dans ce cas, l'agent ne doit pas suivre le principal par obligation mais par intérêt [Mougeot, 1994]. En revanche le remboursement intégral des coûts observés ne serait jamais, quelle que soit l'organisation sanitaire, un système de paiement optimal. Les modèles multi-tâches en santé intègrent le dilemme coût / qualité et la Nature égoïste versus altruiste du producteur de soins.

3.2. Les modèles multi-principaux ou la théorie de l'agence commune

Un agent passe un contrat synchrone avec plusieurs principaux. Autrement dit, un individu (l'agent) choisit une action, qui affecte son utilité et l'utilité des autres joueurs (les principaux) ; ces derniers offrent à l'agent une rétribution monétaire fonction de l'action choisie. Si les principaux coopèrent entre eux, nous retrouvons la configuration de base de la relation d'agence (un agent, un principal). Toutefois, la coopération impose des règles, qui

permettent d'obtenir des engagements irrévocables. Elle reste cependant difficile à établir. Les modèles multi-principaux sont, par conséquent, basés sur des comportements non coopératifs entre les principaux. Certains reposent sur une information complète, sans aléa moral (l'effort fourni par l'agent est observable sans coût par les principaux), ni sélection adverse entre les principaux (chaque principal dispose de la même information que les autres principaux) [Laussel et *al.*, 1996]. D'autres modèles postulent l'aléa moral [Bernheim et *al.*, 1986], où chaque principal demande un effort différent à l'agent. La notion de « principal de principal » [Spiller et *al.*, 1991] est adaptée des travaux de Holmström relatifs aux comportements entre plusieurs agents [Holmström, 1982]. En santé, A. Blomqvist instaure une relation entre l'assureur, le médecin et le patient, chacun détenant sa propre information. Le médecin connaît, contrairement au patient, la Nature exacte de la maladie alors que l'assureur ne dispose bien souvent que d'une information très limitée. Certains systèmes, comme le Health Maintenance Organisation (H.M.O.) dans lequel les médecins deviennent salariés, résolvent les problèmes liés à l'asymétrie de l'information. Le médecin n'est plus incité, contrairement à un système de paiement à l'acte, à facturer davantage les soins [Blomqvist, 1991]. Certaines limites du modèle de A. Blomqvist sont néanmoins formulées par M. Gaynor [Gaynor, 1994].

3.3. Les modèles multi-agents

Ces modèles, dérivés du cadre théorique de l'asymétrie d'information [Laffont et *al.*, 1986], comprennent un principal et plusieurs agents. En situation de monopole naturel, les agents se trouvent en concurrence fictive (chaque offreur se trouve en situation de concurrence fictive car le marché n'existe pas à proprement parlé) ou en concurrence organisée (des quasi-marchés sont organisés par l'Etat). En économie de la santé, ces modèles formalisent les relations entre les autorités sanitaires et les producteurs de soins et examinent les avantages relatifs de la concurrence fictive (tarification à la pathologie) par rapport à la concurrence organisée (achats de soins par la tutelle) [Mougeot et *al.*, 1997].

3.4. Les modèles dynamiques

L'intégration d'une variable temporelle au sein des modèles d'agence impose de prendre en compte certains effets comme la réputation et les aléas de moralité [Rubinstein et *al.*, 1983], le sens de l'altruisme [Barrow, 1996], la notion d'atmosphère [Williamson, 1975] ou la relation de don [Titmuss, 1970].

3.5. Les modèles hiérarchiques ou de délégation

Contraint par l'asymétrie d'information, le principal délègue à un tiers le contrôle de l'agent [Barrow, 1996 ; Strausz, 1997]. Ce tiers, dénommé superviseur ou agence de régulation, est introduit entre le principal et l'agent. Le recours à la délégation est justifié par la présence d'un double risque moral [Macho-Stadler et *al.*, 1991]. Le principal peut directement surveiller l'agent, il peut déléguer cette tâche à l'agent lui-même ou à un superviseur. Si le principal opte pour la première possibilité, il doit fournir un effort de contrôle. Si cet effort n'est pas vérifiable par l'agent, le principal peut mentir à l'agent afin de lui réduire sa rémunération. Nous sommes alors en présence d'un double aléa moral : premièrement, le principal n'observe pas ou mal l'effort de l'agent ; deuxièmement, l'agent ne peut observer l'effort de surveillance du principal. I. Macho-Stadler et J. Perez-Castrillo montrent que déléguer la tâche à un tiers dont l'effort de surveillance est observable par le principal est la solution préférable en présence d'un double aléa moral. La présence d'un superviseur pour exercer le contrôle peut aussi reposer sur l'hypothèse que le principal n'a pas toujours le temps ou les connaissances pour superviser l'agent [Calvo et *al.*, 1978 ; Tirole, 1986]. Dès lors, un risque de collusion apparaît entre, d'une part, le superviseur et l'agent à l'encontre du principal, d'autre part, le principal et le superviseur à l'encontre de l'agent. Certains modèles, comme celui développé par R. Strausz, intègrent ce risque. À la contrainte d'incitation individuelle (stimuler l'effort de l'agent) s'ajoute alors la contrainte de collusion et la difficulté de centraliser les contrats incitatifs dans la hiérarchie [Laffont et *al.*, 1998]. En santé, les autorités n'ont donc pas toujours intérêt à déléguer le contrôle d'un ou plusieurs agents à un superviseur voire davantage, lorsque l'asymétrie d'information repose simultanément sur différentes variables comme l'efficacité et la qualité. Les défauts de coordination entre deux agences de régulation peuvent, en effet, entraîner une surconsommation des ressources [Rizzo et *al.*, 1994].

Nous consacrons la section suivante à l'application de ce dernier type de modèle d'agence aux ordonnances hospitalières de 1996, le Ministre de la santé jouant le rôle de principal, l'A.R.H. et l'hôpital, représentés par leurs directeurs, celui respectivement de superviseur et d'agent.

SECTION 3 : UN MODÈLE D'AGENCE APPLICABLE AUX ORDONNANCES HOSPITALIÈRES DE 1996

Nous développons, dans un premier temps, un modèle fondamental applicable au service public hospitalier avec de l'aléa moral et l'introduction d'un superviseur. Puis, nous appliquons ce modèle au P.M.S.I. et aux ordonnances de 1996 [Perrier et *al.*, 2001].

1. Le modèle fondamental

1.1. Les acteurs du modèle

Considérons un modèle d'agence usuel avec de l'aléa moral et deux joueurs, le principal et l'agent. Le principal rémunère l'agent pour accomplir une tâche et l'agent possède un avantage informationnel sur son type, ses actions ou le monde extérieur en un certain point du jeu. On est en information asymétrique. L'information est complète au sens de E. Rasmusen, en ce sens que la Nature ne joue pas en premier [Rasmusen, 1995]. En outre, cette information est incertaine en ce sens qu'un mouvement de la Nature se produit avant que les mouvements de tous les joueurs soient achevés.

Nous reprenons une partie de la démarche de I. Macho-Stadler et J.D. Pérez-Castrillo [Macho-Stadler et *al.*, 1992]. Situé dans un cadre d'aléa moral, le principal instaure un mécanisme de contrôle pour obtenir de l'information sur le comportement de l'agent. Notons qu'aucune variable vérifiable n'est disponible gratuitement. Il existe une technologie de contrôle et pour la mettre en œuvre, le principal passe un contrat avec un superviseur. Dans ce type de structure, le principal se charge de concevoir un contrat incitatif optimal. L'agent fournit un certain effort et la Nature ajoute du bruit. Le principal paie l'agent selon ses résultats et conserve le cas échéant un surplus. Nous pouvons néanmoins penser, qu'une procédure de contrôle permettrait d'améliorer les choses dans la mesure où elle apporterait de l'information sur le comportement de l'agent [Hart et *al.*, 1987]. La référence au contrôle induit bien souvent l'introduction d'un troisième acteur, le Superviseur. La littérature fournit un certain nombre d'explications pour justifier la présence d'un superviseur [Strausz, 1995a ; Strausz, 1996 ; Demougin et *al.*, 1997 ; Dittmann, 1999]. Par exemple, le principal peut retirer un gain de l'introduction d'une tierce personne ou n'a tout simplement pas le temps de se

consacrer lui-même au contrôle de l'agent.

Le Ministre de la santé, en tant que principal a pour mission d'assurer une production optimale des soins en termes quantitatifs et qualitatifs permettant de préserver ses administrés en bonne santé. Il confie cette tâche, sous contrainte de son enveloppe budgétaire l'O.N.D.A.M., en partie aux agents c'est à dire aux hôpitaux. Le Ministre peut chercher à augmenter son budget pour améliorer sa position et son pouvoir [Tullock, 1978]. Il peut aussi fixer un niveau de production des soins correspondant à la demande et chercher à minimiser son enveloppe budgétaire sous contrainte de ce volume de production. Dans notre modèle, le Ministre rémunère l'agent (l'hôpital) chargé de participer à l'offre de soins sur une zone géographique déterminée et le superviseur sensé contrôler les résultats de l'agent. L'objectif du principal consiste alors à minimiser simultanément la D.G.F. de l'agent et les émoluments du superviseur.

Avant les ordonnances de 1996, les D.D.A.S.S. exerçaient une partie du pouvoir de tutelle sur les établissements de soins par l'intermédiaire notamment de la détermination et de l'approbation de leurs budgets annuels. La mission de la D.D.A.S.S. était complétée par celle des D.R.A.S.S. pour la planification à moyen terme des besoins en équipements. Sans remettre en cause explicitement ces deux structures, les ordonnances de 1996 suppriment de fait leur autorité en matière d'allocation des budgets aux hôpitaux. Cette mission étant détenue depuis par l'A.R.H., nous assimilons donc le superviseur au directeur de l'A.R.H. [Cueille et *al.*, 2000]. L'A.R.H. doit veiller à la satisfaction des objectifs de santé publique à un coût minimal.

L'hôpital, en tant qu'agent, cherche à cumuler les points I.S.A. pour maximiser son budget théorique tout en contrôlant l'évolution de son budget réel. Ainsi, lorsque la valeur du point I.S.A. de l'agent est inférieure à la valeur du point I.S.A. régional, ce dernier peut dénoncer une situation de sous-dotation et réclamer une réévaluation de sa D.G.F. Le superviseur peut aussi considérer que l'agent fournit un effort élevé. L'hôpital dans son rôle d'agent, s'engage à couvrir une partie de la demande de soins de la population en échange du financement. Mais dans quelle mesure respecte-t-il réellement son contrat, sachant que l'asymétrie d'information introduit de l'opacité dans la relation d'agence ?

1.2. Le modèle

La fonction d'utilité de l'agent est additivement séparable entre deux composantes, w (le salaire ou gain) et e (l'effort), ce qui induit implicitement une indépendance d'additivité en

probabilité qui nous éloigne de la réalité. La variable d'effort est censée varier sur un intervalle $[0,1]$. La fonction d'utilité s'écrit

$$U(w, e) = u(w) - v(e) \quad (1)$$

Cette hypothèse de séparabilité signifie que l'aversion de l'agent vis à vis du risque ne varie pas en fonction de l'effort qu'il fournit. Nous posons les hypothèses usuelles sur les fonctions u et v . La fonction u est croissante et concave (au sens large)

$$u'(w) > 0 \text{ et } u''(w) \leq 0$$

Ainsi l'agent peut être, soit neutre vis-à-vis du risque, soit risquophobe. Par ailleurs, la désutilité de l'agent croît avec l'effort, ainsi

$$v'(e) > 0 \text{ et } v''(e) \geq 0$$

En outre, nous posons que

$$v'(0) = 0 \text{ et } v'(1) = +\infty$$

L'utilité de réservation de l'agent, celle offerte par des opportunités externes est notée \underline{U} . Pour l'hôpital, l'utilité de réservation pourrait correspondre à ne pas prendre en charge un patient dont le coût serait indubitablement supérieur au coût médian du G.H.M. Ce dernier reste toutefois contraint par sa mission de service public.

Le principal ou un superviseur peut exercer le contrôle. Nous supposons dès le départ, que le contrôle sera exercé par le superviseur, le principal n'ayant pas le temps de superviser l'agent. La fonction objectif du superviseur est identique à ce que serait celle du principal, si ce dernier exerçait le contrôle. La technologie de surveillance reste aussi la même, que ce soit l'un ou l'autre, qui exerce le contrôle. A cet égard, toute distorsion au profit du superviseur remettrait en cause le schéma de délégation. Dans cette optique l'utilité de réservation de l'éventuel superviseur est $\underline{V} = 0$. Ainsi, nous écartons les effets dus à un éventuel avantage technique du superviseur.

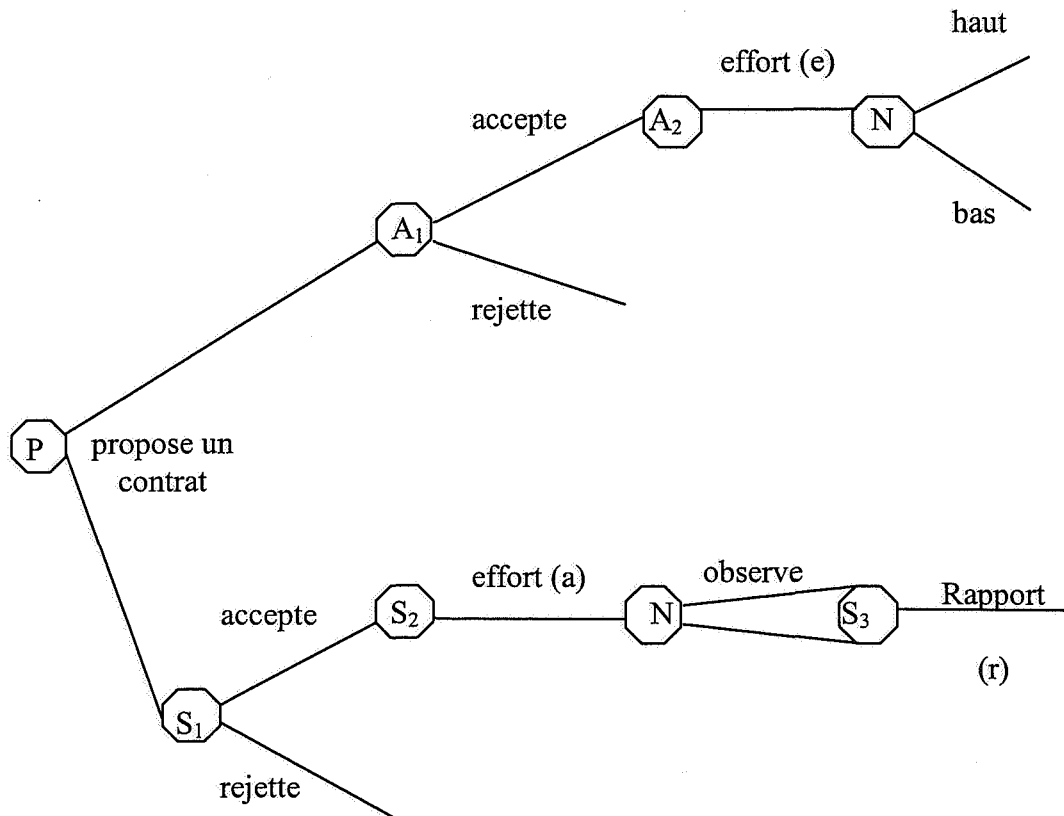
La fonction objectif du principal est une fonction de coûts à minimiser. Elle intègre les paiements aux agents engagés.

Nous avons

$$C(y) = y \quad (2)$$

où y représente les versements monétaires. Compte tenu des contraintes budgétaires, nous pouvons effectivement penser, que l'objectif du principal est de minimiser à la fois la D.G.F. de l'agent et la rémunération du superviseur. Le déroulement du jeu peut se résumer dans la figure suivante :

Figure 11
Le déroulement du jeu



Avec P : le principal, S : le superviseur et A : l'agent.

Dans un premier temps, le principal décide d'embaucher un superviseur. Il offre un contrat vecteur w à l'agent et un contrat vecteur s au superviseur. Il veut obtenir de l'agent un niveau d'effort donné e^* . Pour ce faire, il choisit les contrats, qui minimisent les paiements espérés de l'agent et du superviseur. La fonction objectif de l'A.R.H. peut donc s'écrire

$$\Phi(s, a) = s - d(a) \quad (3)$$

où s représente les versements monétaires, $a \in [0,1]$ est l'effort fourni pour le contrôle, $d(a)$ est la fonction de désutilité telle que

$$d'(a) > 0 \quad \text{et} \quad d''(a) \geq 0$$

$$d'(0) = 1 \quad \text{et} \quad \lim_{a \rightarrow 1} d'(a) = +\infty$$

Si l'hôpital accepte le contrat, il fournit un effort e dont le résultat n'est pas vérifiable. Nous supposons, que seuls deux résultats sont possibles : Haut = H (ou Bon) et Bas = B (ou Mauvais). La Nature détermine l'état des choses, c'est-à-dire que la probabilité d'obtenir H est $p(e)$ et celle d'obtenir B vaut $[1-p(e)]$. Nous supposons que la fonction p est croissante et concave.

$$p'(e) > 0 \quad \text{et} \quad p''(e) < 0$$

Si le superviseur accepte le contrat, il fournit un effort a . Dans la mesure où la Nature ajoute du bruit, nous avons une probabilité $q(a)$ d'observer sans erreur l'effort e et une probabilité $[1-q(a)]$ de ne pas pouvoir fournir les preuves de l'effort (le superviseur n'observe rien). Nous supposons que

$$q'(a) > 0 \quad \text{et} \quad q''(a) < 0$$

La technologie de contrôle est telle que l'on ne peut avoir que trois signaux possibles : ne rien observer (R), observer H ou observer B . L'ensemble d'observation est alors

$$O = \{R, H, B\}$$

La Nature ayant révélé l'action de l'hôpital, l'A.R.H. décide soit de rendre publique l'information obtenue, soit de la retenir. Dans les faits, le superviseur transmet un rapport (r), qui peut taire l'information sans jamais donner une information fausse. Formellement, nous avons

$$r = \{\emptyset, O\}$$

Nous faisons l'hypothèse que lorsque le superviseur assure le contrôle, le principal ne peut distinguer entre R et \emptyset , ce qui se traduit formellement par

$$R \equiv \emptyset$$

Au total, le contrat de l'agent est un vecteur w , tel que

$$w = (w_R, w_i) \quad i = H, B \quad (4)$$

où w_R représente le salaire versé si le rapport ne dit rien et w_i le salaire versé en fonction du résultat.

De même, le salaire versé au superviseur (les valeurs possibles étant une fonction du rapport rendu) est un vecteur s tel que

$$s = (s_R, s_i) \quad i = H, B \quad (5)$$

Le superviseur ne révélera l'information obtenue, que si $s_R < s_i$, $i = H, B$. Dès lors, le Ministre de la Santé peut verser à l'A.R.H. un paiement uniforme σ si cette dernière observe un résultat et lui verser s_R si elle n'observe rien. Le contrat du superviseur se ramène alors au vecteur

$$s = (s_R, \sigma) \quad (6)$$

Le paiement est fonction de la non observation ou de l'observation d'un résultat, quel qu'il soit. Par ailleurs, si $s_R < \sigma$, le Superviseur se voit contraint de transmettre l'information dont il dispose au principal, information nécessaire à ce dernier.

La « complication » du modèle est possible en introduisant une probabilité de contrôle du superviseur [Strausz, 1995b]. En termes de jeux, l'agent et le superviseur joueraient alors à un moment donné un jeu à mouvement simultané : l'hôpital choisissant son effort et l'A.R.H. choisissant de contrôler ou non.

La recherche de contrats optimaux s'effectue usuellement en utilisant une approche développée par S. Grossman et O. Hart [Grossman et al., 1983], qualifiée par D. Fudenberg et J. Tirole de procédure en 3 étapes [Fudenberg et al., 1990]. Ainsi, nous débutons l'analyse en raisonnant à couples d'efforts donnés (e^*, a^*) .

Analysons maintenant ce qu'advient les contraintes de l'agent et du superviseur. La contrainte de participation de l'agent le conduit à accepter effectivement le contrat, si son espérance d'utilité est supérieure ou égale à celle qu'il pourrait atteindre hors de la relation contractuelle (utilité de réservation). Elle s'énonce comme suit

$$q(a^*) \cdot [p(e^*)u(w_H) + (1 - p(e^*))u(w_B)] + [1 - q(a^*)]u(w_R) - v(e^*) \geq \underline{U} \quad (7)$$

La contrainte incitative reflète le fait, que l'agent choisira l'effort qui maximise sa fonction objectif, soit³²

$$e^* \in \arg \operatorname{Max}_e \left\{ \begin{array}{l} q(a^*) \cdot [p(e^*)u(w_H) + (1 - p(e^*))u(w_B)] \\ + [1 - q(a^*)]u(w_R) - v(e^*) \end{array} \right\} \quad (8)$$

Supposer que l'effort de l'agent soit une variable continue, $e \in [0,1]$, entraîne *in fine* un problème de double "maximisation". C'est la raison pour laquelle, nous substituons à ce problème de maximisation de l'agent sa condition de premier ordre grâce à l'approche de premier ordre, [Holmström, 1979]. Ainsi, nous remplaçons (8) par

$$q(a^*) \cdot [u(w_H) - u(w_B)] = \frac{v'(e^*)}{p'(e^*)} \quad (9)$$

Concernant le superviseur, comme le contrat est $s = (s_R, \sigma)$, sa contrainte de participation s'écrit

$$q(a^*) \cdot \sigma + [1 - q(a^*)] \cdot s_R - d(a^*) \geq \underline{V} = 0 \quad (10)$$

et sa contrainte incitative qui l'amène à choisir l'effort a^* est

$$a^* \in \arg \operatorname{Max}_a \{ q(a^*) \cdot \sigma + [1 - q(a^*)] \cdot s_R - d(a^*) \} \quad (11)$$

L'incitation du directeur de l'A.R.H. à exercer correctement sa mission est forte, puisque ce dernier, comme nous l'avons vu, est nommé en Conseil des Ministres. En conséquence, la garantie de son emploi est de sept jours et celle de son salaire de six mois. De plus,

³² Pour tout ensemble Z et toute fonction $f : Z \rightarrow R$, $\arg \max_{y \in Z} f(y)$ correspond à l'ensemble des points de Z qui maximise la fonction f , soit $\arg \max_{y \in Z} f(y) = \{y \in Z / f(y) = \max_{z \in Z} f(z)\}$ [Myerson, 1997 ; p. 53].

l'hypothèse, que son utilité de réservation V soit nulle, est une quasi-obligation. Comme le superviseur est contraint de se situer dans le cadre contractuel, sa contrainte de participation (10) est toujours satisfaite. Enfin $q(a)$ est vraisemblablement proche de 1, puisque l'A.R.H. contrôle l'agent sous forme notamment d'audits complets, d'enquêtes ponctuelles d'une activité, avec un accès au dossier médical du patient.³³

Nous retenons à nouveau l'approche du premier ordre et nous substituons à (11)

$$q(a^*) \cdot [\sigma - s_R] = d'(a^*) \quad (12)$$

Cela étant, si le principal engage le superviseur, il doit choisir les contrats w et s , qui sont solutions du programme de minimisation de coûts

$$\underset{(w,s)}{\text{Min}} \left\{ \begin{array}{l} q(a^*) \cdot [p(e^*) w_H - (1 - p(e^*)) w_B] + [1 - q(a^*)] \cdot w_R \\ + q(a^*) \cdot \sigma + [1 - q(a^*)] \cdot s_R \end{array} \right\} \quad (\text{Pg1})$$

Pour a^* donné, la contrainte de participation de S est saturée dès que

$$q(a^*) \cdot \sigma + [1 - q(a^*)] \cdot s_R = d(a^*)$$

Le coût du superviseur est alors égal à la désutilité de son propre effort. Le programme (Pg1) est donc équivalent au programme

$$\underset{(w)}{\text{Min}} \left\{ \begin{array}{l} q(a^*) \cdot [p(e^*) w_H - (1 - p(e^*)) w_B] \\ + [1 - q(a^*)] \cdot w_R + d(a^*) \end{array} \right\} \quad (\text{Pg2})$$

Notons γ et μ les multiplicateurs associés respectivement à la contrainte de participation et à la contrainte incitative et $L(w, \gamma, \mu)$ le Lagrangien. Les conditions du premier ordre sont

$$\frac{\partial L}{\partial w_R} = 0 \Rightarrow \lambda = \frac{1}{u'(w_R)} > 0 \quad (\text{C1})$$

³³ Se référer au paragraphe 3.4. « Les contrôles de l'A.R.H. et des hôpitaux », p. 40.

$$\frac{\partial L}{\partial w_H} = 0 \Rightarrow \mu = p(e^*) \cdot \left[\frac{1}{u'(w_H)} - \frac{1}{u'(w_R)} \right] \quad (C2)$$

et enfin

$$\frac{\partial L}{\partial w_B} = 0 \Rightarrow [1 - p(e^*)] \cdot [1 - \lambda u'(w_b)] + \mu u'(w_b) = 0$$

Soit, en utilisant (C1) et (C2), une relation dont la solution en $u'(w_R)$ est donnée par

$$\frac{1}{u'(w_R)} = \frac{1 - p(e^*)}{u'(w_B)} + \frac{p(e^*)}{u'(w_H)} \quad (C3)$$

Cette condition est indépendante de l'effort de contrôle exercé par le superviseur (a^*). Elle montre, que l'inverse de $u'(w_R)$ est égal à la moyenne harmonique des utilités marginales $u'(w_B)$ et $u'(w_H)$ pondérées par les probabilités d'obtenir les résultats correspondants.

En combinant (C2) et (C3), il est possible d'ordonner les salaires constitutifs du contrat de A. Nous obtenons alors

$$w_B < w_R < w_H \quad (13)$$

C'est-à-dire qu'en l'absence de délégation du rôle de supervision, on a nécessairement

$$w_R \geq w_i, \quad i = H, B$$

La démonstration de (13) est la suivante : selon (C2) et les conditions de Kuhn et Tucker nous avons

$$\left[\frac{1}{u'(w_H)} - \frac{1}{u'(w_R)} \right] > 0 \Rightarrow u'(w_R) > u'(w_H)$$

Comme u est croissante et concave, nécessairement $w_H > w_R$

Par ailleurs, compte tenu de (C3)

$$\frac{1-p(e^*)}{u'(w_B)} + \frac{p(e^*)}{u'(w_H)} < \frac{1}{u'(w_H)}$$

d'où

$$\frac{1-p(e^*)}{u'(w_B)} < \frac{1-p(e^*)}{u'(w_H)} \Leftrightarrow u'(w_B) > u'(w_H)$$

et donc pour les même raisons, $w_B < w_H$

Il reste à classer w_B et w_R . Une simple transformation de (C3) donne

$$\frac{1-p(e^*)}{u'(w_R)} + \frac{p(e^*)}{u'(w_R)} = \frac{1-p(e^*)}{u'(w_B)} + \frac{p(e^*)}{u'(w_H)}$$

Compte tenu de (C2) et des conditions de Kuhn et Tucker, nous avons

$$\frac{p(e^*)}{u'(w_H)} > \frac{p(e^*)}{u'(w_R)} \Rightarrow \frac{1-p(e^*)}{u'(w_B)} + \frac{p(e^*)}{u'(w_H)} > \frac{1-p(e^*)}{u'(w_B)} + \frac{p(e^*)}{u'(w_R)}$$

d'où

$$\begin{aligned} \frac{1-p(e^*)}{u'(w_R)} + \frac{p(e^*)}{u'(w_R)} &> \frac{1-p(e^*)}{u'(w_B)} + \frac{p(e^*)}{u'(w_R)} \\ \Leftrightarrow \frac{1-p(e^*)}{u'(w_R)} &> \frac{1-p(e^*)}{u'(w_B)} &\Leftrightarrow u'(w_R) < u'(w_B) \end{aligned}$$

et compte tenu des propriétés de u , $w_B < w_R$.

Toutefois, nous montrons dans le paragraphe suivant que la qualification du résultat de l'agent s'avère, compte tenu des spécificités du marché de la santé, bien souvent délicate.

2. Sélection adverse, risque moral et qualification du résultat de l'hôpital

2.1. Le problème de la qualification du résultat de l'agent

L'application du modèle au service public hospitalier pose le problème majeur de la qualification du résultat de l'agent. En d'autres termes, que signifie réellement la typologie « résultat élevé (*H*) » versus « résultat faible (*B*) » ? Ainsi, lorsque la valeur du point I.S.A. de l'agent est inférieure à la valeur du point I.S.A. régional, le superviseur peut considérer que l'agent fournit un effort élevé sauf à considérer qu'il existe d'autres types de mesure permettant d'affiner la typologie. Mais le résultat de l'agent peut être élevé en termes qualitatifs (confort du patient, performance des traitements, etc...) et faible en termes de coûts (valeur élevée du point I.S.A.). Une guérison ne signifie pas nécessairement que les soins prodigués sont bons, et l'aggravation de la maladie que les soins sont mauvais. En conséquence, il est nécessaire de relativiser la productivité mesurée par les coûts d'un établissement de santé grâce à des critères d'efficacité comme le taux de survie ou les infections nosocomiales.

2.2. Sélection adverse et risque moral dans le service public hospitalier

M.V. Pauly montrait déjà, à la fin des années 1960, que la demande plus importante de soins observée chez les patients ayant souscrit une assurance de santé ne résultait pas d'une attitude perfide de ce dernier, mais d'un comportement économique rationnel [Pauly, 1968]. Les propos de M.V. Pauly étaient alors discutés par K.J. Arrow, qui montrait les qualités de son approche en terme d'aléa moral [Arrow, 1970]. Deux décennies plus tard, les questions relatives au risque moral demeurent au cœur de notre analyse. En effet, les réformes de 1996 réduisent le risque moral dans le sens où le superviseur observe mieux la production de l'agent à travers les bases de R.S.A. et la production des I.S.A. Toutefois, les résultats de l'hôpital en termes de coûts et de qualité restent difficiles à appréhender compte tenu de la nature du produit et de la non incitation du mode de financement du service public hospitalier à chiffrer les coûts. En dehors des études médico-économiques ponctuelles, les coûts ne sont pas systématiquement évalués. Le passage à un superviseur unique concernant l'attribution de la D.G.F. limite la possibilité pour l'agent de tirer parti de l'imbrication des responsabilités et des prises de décisions croisées [Moison et *al.*, 1996]. Ainsi, la transparence des différents financements alloués est meilleure.

Les ordonnances hospitalières de 1996 atténuent l'anti-sélection car dans le cadre du contrat d'objectifs et de moyens, les caractéristiques et les objectifs de l'agent sont détaillés avant la signature du contrat. La procédure d'accréditation de l'A.N.A.E.S. permet aussi de réduire l'anti-sélection, puisque certains financements spécifiques ne sont accordés que si l'agent offre les caractéristiques souhaitées en terme de savoir faire, de qualité, etc... Un superviseur externe comme l'A.N.A.E.S. offre une meilleure observabilité des résultats, si nous supposons une absence de collusion entre le superviseur et l'agent.

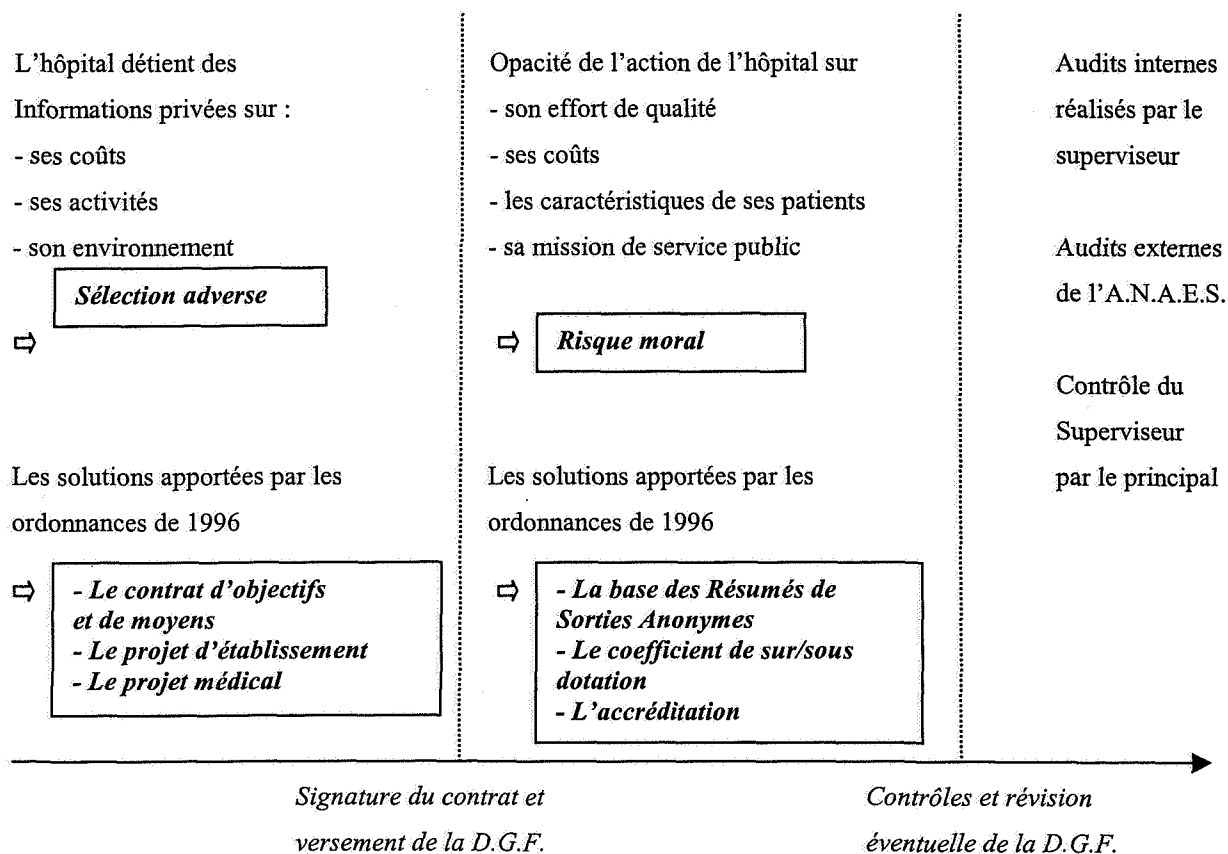
Bien que les réformes actuelles réduisent le risque moral (le superviseur observe mieux la production de l'agent après le financement) et la sélection adverse (le principal connaît mieux les caractéristiques et les objectifs de l'agent avant le financement), l'agent peut toujours tricher [Cueille et *al.*, 2000]. L'hôpital peut sélectionner les patients dont le coût probable du traitement reste inférieur à la rémunération théorique du G.H.M. Il peut également gonfler artificiellement le nombre de points I.S.A. en re-hospitalisant le patient. L'hôpital peut aussi affecter des patients dans un G.H.M., qui génère plus de points I.S.A. Par exemple, avant la sixième version des G.H.M. mise en place en 2000, l'hôpital devait normalement affecter un patient ayant un Prélèvement de Cellules Souches du Sang Périphérique (P.C.S.S.P.) dans le G.H.M. 827 intitulé « *Greffes d'organes : ambulatoire, sans acte opératoire* », qui générait 352 points I.S.A., soit une dotation théorique de 4491 francs (684,7 €).³⁴ Pour le même acte thérapeutique, l'hôpital pouvait hospitaliser le patient au moins 24 heures et, par là même, le basculer dans le G.H.M. 584 intitulé « *Greffe de moelle* », qui génère 11567 point I.S.A soit 147595 francs (22500,7 €). Enfin, si la qualité des soins est difficilement observable, l'hôpital n'est pas incité, dans ce contexte de rémunération, à l'améliorer. Nous retrouvons ici d'autres limites inhérentes au P.M.S.I. De plus, comment déterminer un mécanisme de révélation des coûts dans le service public hospitalier ? Si un hôpital, dont les coûts sont faibles, déclare des coûts élevés pour soutirer une rente au principal, il risque un audit du superviseur pour analyser la nature du surcoût. Dans un système qui devient concurrentiel, il risque même une révision de son financement au profit d'un autre hôpital plus « compétitif », voire à l'extrême une fusion que peut lui imposer le Directeur de l'A.R.H. Si l'hôpital annonce des coûts plus faibles, il risque de se voir diminuer sa D.G.F.

L'apport des ordonnances hospitalières de 1996 par rapport au risque moral et à la sélection adverse peut être synthétisé à partir du schéma suivant.

³⁴ En prenant la valeur du point I.S.A. de la région Rhône-Alpes en 1997, soit 12,76 francs (1,95 €).

Figure 12

Synthèse de l'asymétrie d'information dans le service public hospitalier



Source : Figure inspirée des travaux de S. Cueille et A. Renucci [Cueille et al., 2000].

Compte tenu, d'une part des négociations croissantes avec le superviseur, d'autre part des difficultés de qualification des résultats, l'agent recourt de plus en plus aux évaluations de coûts pour justifier sa D.G.F.

3. Le rôle croissant des évaluations de coûts en santé dans les négociations entre l'agent et le superviseur : l'exemple de la revalorisation du G.H.M. 827.

L'agent constate que le P.M.S.I. rémunère insuffisamment les soins ambulatoires en cancérologie. Une évaluation médico-économique d'un soin ambulatoire sert alors à

démontrer au principal l'insuffisante rémunération théorique. Nous présentons, à titre d'illustration, l'évaluation que nous avons conduite au sein de l'unité des Prélèvements de Cellules Souches du Sang Périphérique (P.C.S.S.P.) du C.L.B. à partir des données de 1997 [Perrier et *al.*, 2000].

3.1. Le Prélèvement de Cellules Souches du Sang Périphérique

D'un point de vue médical, le P.C.S.S.P. constitue une étape des intensifications thérapeutiques avec greffes de cellules souches du sang périphérique pratiquées auprès de patients atteints de tumeurs malignes sensibles à la chimiothérapie. Les intensifications thérapeutiques débutent par une phase de mobilisation avec chimiothérapie au décours de laquelle le patient reçoit des facteurs de croissance hématopoïétiques, qui stimulent les cellules souches (cellules CD34⁺) et permettent ainsi leur augmentation dans le sang périphérique [Hass et *al.*, 1994]. Ces cellules sont alors prélevées (P.C.S.S.P.) et congelées par le laboratoire d'analyse et de congélation des greffons. Elles sont décongelées et réinjectées (greffe) après une chimiothérapie intensive (phase de conditionnement). Le patient reste hospitalisé en unité stérile tant qu'il est en aplasie (taux de plaquettes et de polynucléaires neutrophiles insuffisants). En 1997, l'activité s'est élevée à 323 P.C.S.S.P., contre 484 en 1996 et 527 en 1995. Les autogreffes ont été de 160 en 1997, contre 171 en 1996 et 147 en 1995. Suite à une meilleure qualité des greffons, le nombre moyen de P.C.S.S.P. par patient est passé de 3,25 en 1995 à 2,06 en 1997. Les diagnostics les plus fréquents des patients greffés en 1997 étaient les lymphomes (23%), les tumeurs du sein (26%), les myélomes (16%), les sarcomes d'Ewing (5%), les neuroblastomes (4%). Cette unité comprend un médecin spécialiste à mi-temps et son secrétariat, deux infirmières à temps plein qui contrôlent deux appareils à P.C.S.S.P. Un cadre à temps plein est responsable du laboratoire d'analyse et de congélation des greffons ; il est accompagné de deux techniciens de laboratoire à temps plein. La durée d'amortissement des immobilisations est de 5 ans (conteneurs, hottes, étuves centrifugeuses...), sauf la soudeuse et les deux appareils de P.C.S.S.P. amortis sur 8 ans.

Une observation détaillée sur le site et de nombreux entretiens avec le personnel nous ont permis de décomposer un P.C.S.S.P. en quatre étapes :

(1) La préparation au P.C.S.S.P. : elle est principalement réalisée par une infirmière et comprend la préparation de la machine (mise en place du kit et des poches, programmation), la préparation du plateau de soins et du petit matériel (aiguilles et compresses stériles,

seringues, etc...), l'accueil et la préparation du patient (nettoyage du cathéter, prise de sang, etc) et la consultation médicale (effectuée par le médecin).

(2) Le P.C.S.S.P. : il inclut la surveillance du patient par le médecin et les infirmières, les injections éventuelles de calcium, les démarches administratives.

(3) La fin du P.C.S.S.P. : réalisée par une infirmière, elle intègre la restitution sanguine, les derniers actes médicaux (pouls, tension artérielle, prise de sang, nettoyage du cathéter...), l'évacuation des éléments jetables (déprogrammation de la machine, retrait du kit usagé, nettoyage), la désinfection du plateau de soins et des surfaces.

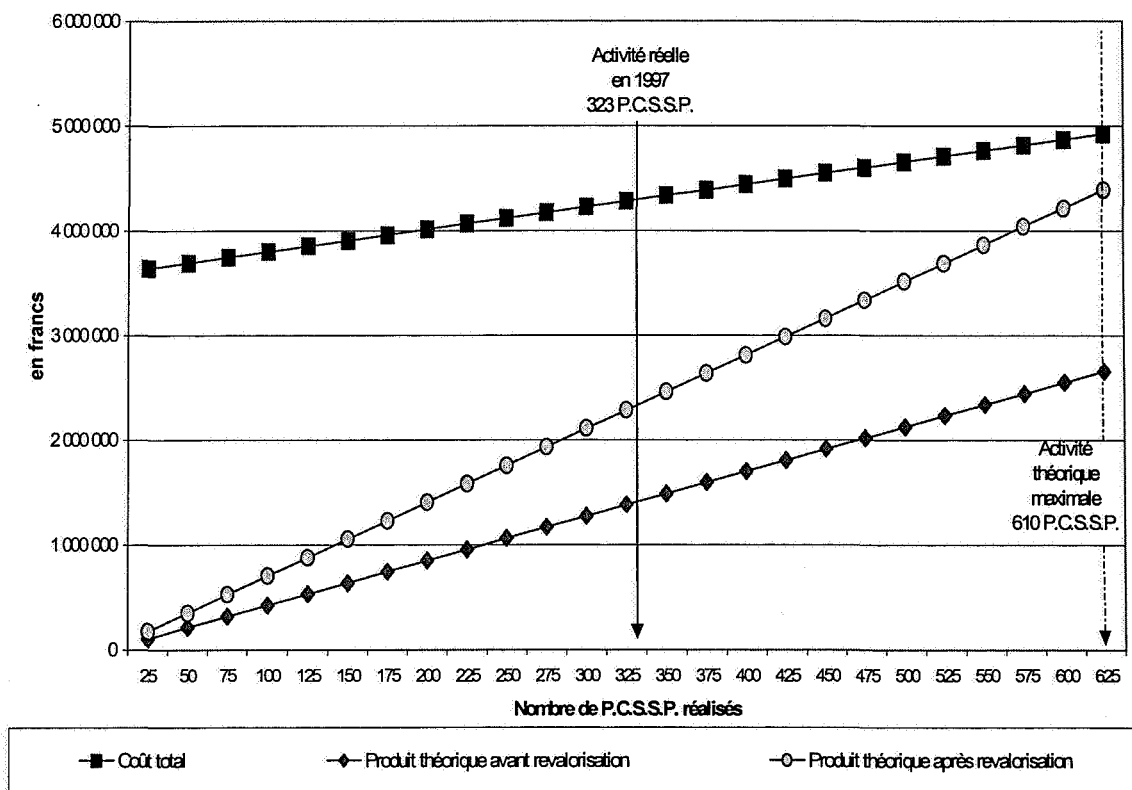
(4) L'analyse et la congélation des greffons : elles sont réalisées par les techniciens de laboratoire et concernent la préparation à la congélation et la congélation elle-même. Cette préparation comprend entre autres la numération des cellules souches recueillies, l'élaboration des poches de congélation avec l'adjonction de cryoprotecteur (la congélation dans l'azote liquide ferait éclater les cellules), la préparation des tubes d'échantillons, les démarches administratives. L'analyse qualitative et quantitative du greffon conduit soit à la poursuite des prélèvements, soit à leur arrêt. Enfin, la décongélation et la réinjection du greffon au patient nécessitent en particulier le transfert des cellules souches de la poche de congélation vers une poche de perfusion, le retrait du cryoprotecteur.

3.2. Les valorisations monétaires d'un P.C.S.S.P.

Le coût moyen réel d'un P.C.S.S.P. s'est élevé à 13239 francs (2018,3 €) pour 323 P.C.S.S.P. pratiqués en 1997. Le coût moyen optimal aurait été de 8017 francs (1222,2 €), la capacité maximale annuelle de production de l'unité étant estimée à 610 P.C.S.S.P. Le coût marginal jusqu'au 610^{ème} P.C.S.S.P. s'est établi à 4987 francs (760,3 €) et le coût P.M.S.I. à 11838 francs (1804,7 €) sur la base d'une activité annuelle de 323 P.C.S.S.P. Le coût P.M.S.I. proposé par la mission P.M.S.I. [Ministère de l'Emploi et de la Solidarité, 1997], comprend quatre sections d'imputation. La Section d'Imputation 1 (S.I. 1) évalue les courts séjours en unité médicale M.C.O. et la S.I. 6 les supports médico-techniques (laboratoires d'analyses, etc...). La S.I. 4 évalue les activités de blanchisserie et de restauration. Enfin, la S.I. 5 comprend les coûts liés à la logistique et l'administration. Après l'agrégation des coûts de chaque section d'imputation, les Centres Hospitalo-Universitaires (C.H.U.) et les Centres Régionaux de Lutte Contre le Cancer (C.R.L.C.C.) peuvent retrancher un forfait de 13 pour cent correspondant à l'activité de recherche. Sur la base d'une évaluation interne corroborée par d'autres résultats [Smith et *al.*, 1997 ; Woronoff-Lemsi et *al.*, 1997], la revalorisation

du G.H.M. 827 « Greffes d'organes : ambulatoire, sans acte opératoire » s'avérait nécessaire. La dernière version de la classification en G.H.M. [Ministère de l'Emploi et de la Solidarité, 2000], rétroactive au premier janvier 1999, revalorise le G.H.M. 827 de 352 points I.S.A. à 582 points I.S.A. La rémunération théorique de l'hôpital passe alors de 4248,68 francs (647,7 €) à 7024,74 francs (1070,9 €) par P.C.S.S.P. Malgré cette revalorisation, le seuil de rentabilité de l'unité reste supérieur à son volume optimal de production. Il atteint encore 734 P.C.S.S.P. annuels, contre 1701 auparavant.

Figure 13
Les effets de la revalorisation du G.H.M. 827 sur le seuil de rentabilité de l'unité de production

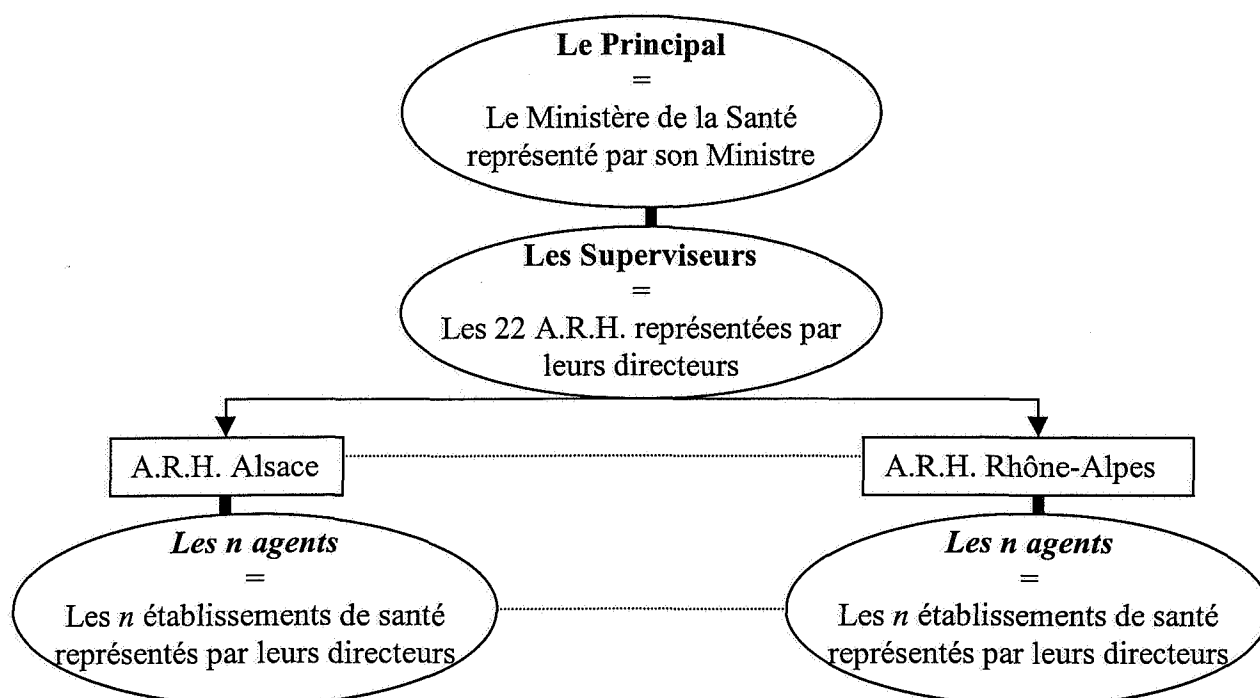


Nous n'avons pas abordé la question importante des sanctions du superviseur en cas de fraude. Quelle attitude doit-il adopter en cas de tricherie ? Combien peut-il ponctionner sur le budget de l'hôpital pour être dissuasif sans mettre en péril la mission de service public du « tricheur » ? Si les obligations du contrat d'objectifs et de moyens ne sont pas remplies, les ordonnances de 1996 sont floues sur les sanctions encourues, même si le directeur de l'hôpital engage sa responsabilité lors de la signature du contrat [Jean, 1997].

CONCLUSION

Ce premier chapitre nous a permis d'appréhender, à partir de la théorie économique, en l'occurrence la relation d'agence, l'organisation et le mode de financement actuels du service public hospitalier. Compte tenu d'une modélisation qui se complexifie rapidement, le modèle présenté simplifie nettement la réalité, puisque nous devrions avoir un modèle hiérarchique simultanément multi-superviseurs et multi-agents construit sur la base de la figure ci-dessous.

Figure 14
Les acteurs du système de santé dans notre relation d'agence



La simplification du modèle présenté semble toutefois acceptable dans la mesure où, lorsque nous aurons introduit précisément la notion de trajectoire optimale du patient (chapitre 3), nous obtiendrons une réalité plus simple avec un principal, un superviseur et un agent unique. Si les ordonnances de 1996 et le P.M.S.I. atténuent l'anti-sélection et le risque moral, le fait de calculer les dotations théoriques à partir des séjours présente, comme nous l'avons vu, des limites. Pour les dépasser, nous développons le concept de la trajectoire du patient. Après avoir défini ce terme, nous montrons, sur le plan théorique, comment les trajectoires des patients peuvent servir d'outil de planification pour le principal, d'instrument de contrôle pour le superviseur et de support de négociation pour l'agent (chapitre 2). Puis, nous appliquons, dans le cadre de notre relation d'agence, le concept de trajectoire à l'oncologie pédiatrique (chapitre 3).