

LA THEORIE DES COÛTS DE TRANSACTIONS

La théorie des coûts de transaction s'est construite en plusieurs temps. Elle repose, dans ses développements actuels, sur certaines hypothèses différentes de l'axiomatique néo-classique, principalement une hypothèse de rationalité différente. La notion de coût de transaction est introduite pour la première fois par Ronald Coase en 1937. Elle ne suscitera aucun commentaire pendant de nombreuses années, avant de valoir à son auteur une gloire tardive (Prix Nobel). C'est grâce aux travaux de Williamson, que la notion de coût de transaction prend toute son ampleur. Grâce à ce concept, la théorie des coûts de transaction est en mesure de rendre compte de l'existence de la firme dans une économie de marché. En outre, elle permet de comprendre les formes de **gouvernance** des transactions. Ainsi, elle permet d'étudier les formes organisationnelles.

Nous présenterons donc les travaux pionniers de Coase à la base de cette théorie. Nous porterons ensuite notre attention sur les travaux de Simon qui permettent de rendre compte de l'existence de la relation d'autorité. Enfin, nous présenterons les travaux d'O.E. Williamson qui constituent à proprement parler la théorie des coûts de transaction.

1) Le coût d'un système de prix : l'apport de R. Coase¹.

Contrairement à la théorie des droits de propriété ou à la théorie de l'agence, la théorie des coûts de transaction n'a pas pour ambition de démontrer la supériorité de la firme capitaliste sur tout autre forme d'organisation de la production. La question à laquelle Coase entend répondre est celle de l'existence de la firme. Pourquoi dans une économie de marché existent-ils des entreprises ?

Pour tenter d'expliquer ce mode de coordination de la production, Coase introduit une notion qui posera les fondements des sciences de l'organisation : l'efficacité organisationnelle. Une forme d'organisation économique est efficace lorsqu'elle minimise les coûts de fonctionnement. Ces coûts de fonctionnement sont appelés les coûts de transaction. En effet, l'unité d'analyse est la transaction. Celle-ci fait toujours l'objet d'un contrat dont la nature caractérise une forme spécifique d'organisation.

Pour Coase, l'entreprise se définit comme une relation d'autorité. L'entreprise met en jeu des contrats asymétriques dans la mesure où les employés acceptent librement la subordination à la volonté de l'employeur. Cette relation s'oppose aux relations sur un marché. Dans un tel mode de coordination, les agents sont réputés symétriques, il n'y a pas de rapport de pouvoir ni de subordination.

La nature de la firme.

La firme est définie comme « *un système de rapports qui apparaît lorsque la direction des ressources dépend d'un entrepreneur [c'est-à-dire dépend de] la ou les*

¹ Coase R.H., [1937], « *The nature of the firm* », *Economica*, Vol. 16, pp. 331-351.

personnes qui, dans un système concurrentiel, prennent la place du mécanisme des prix dans la direction des ressources. »² Cette définition ne dit rien sur la nature de la firme. Coase la définit plus précisément par l'existence de la relation salariale. « *le contrat [de travail] est un contrat par lequel le facteur travail accepte, pour une certaine rémunération, qui peut être fixe ou variable, d'obéir aux directives d'un entrepreneur dans certaines limites.* »³ L'essence de la firme capitaliste repose sur l'existence de cette relation de subordination. Le rapport salarial est spécifique à l'organisation capitaliste de la production : il implique un rapport d'autorité, qu'elle soit directement assumée par le dirigeant ou déléguée à ses représentants (mandataires). Le travail en tant que cession temporaire de la force de travail à l'employeur suppose en effet, le droit pour ce dernier d'en disposer à sa convenance, dans le cadre des termes contractuels et sous réserve de la législation en vigueur. Après avoir précisé ce qui caractérise l'organisation de la production au sein de la firme, Coase tente de répondre à la question de l'émergence de la firme : « *pourquoi la firme émerge-t-elle dans une économie spécialisée dans l'échange?* »⁴

Pourquoi la firme ?

Telle que formulée par Coase, la question présuppose l'existence de marchés sur lesquels interviennent des agents économiques indépendants. Ce type de système économique ne permet pas d'expliquer l'existence de la firme. En effet, l'allocation des ressources se fait sans rapports de force, ni de subordination, les individus libres et autonomes échangent des ressources grâce à l'existence d'un mécanisme de prix.

Dans la firme, l'allocation des ressources ne se fait pas à l'aide du mécanisme de prix, mais par l'autorité. Quelles sont donc les raisons de l'existence d'un tel mode de coordination ? Coase examine plusieurs raisons pour les rejeter et proposer une explication qui tient à la nature coûteuse des transactions.

Coase analyse et critique l'explication de l'existence de la firme avancée par F. Knight⁵. Pour Knight, la firme a une double origine. Son existence est due à l'incertitude qui caractérise la société réelle et a des profils psychologiques des agents. Introduire l'incertitude dans la théorie économique modifie considérablement une économie de marché où tous les demandes sont prévisibles. Les demandes sont inconnues *a priori*. La question devient alors que produire et comment produire ? Prévoir le désir des consommateurs est celle de l'entrepreneur qui se « spécialise » dans les tâches de conception. Ainsi les tâches de conception se séparent-elles des tâches d'exécution. L'incertitude est donc à l'origine de l'existence de la firme (i.e. de l'entrepreneur) et de l'organisation hiérarchique.

F. Knight distingue ensuite deux classes d'agents : les « *plus confiants en eux-mêmes et les plus aventureux assument les risques ou rassurent les timorés et les timides en leur garantissant un revenu spécifié en contrepartie de résultats concrets.* »⁶. Ainsi, à partir de cette hypothèse Knight est-il en mesure d'expliquer

² Coase R.H., [1937], *op. cit.*, p.339.

³Coase R.H., [1937], *op. cit.*, p.337.

⁴ *ibid.*, p. 335.

⁵ Knight F., [1921], « *Risk, uncertainty, and profit* », Harper Torchbooks, New York, (1965).

⁶ *ibid.*, p. 268-269.

l'existence du salariat. Mais cette hypothèse est extrêmement discutable dans la mesure où, s'il s'agit d'un trait psychologique lié à la nature humaine, pourquoi le salariat n'a-t-il pas toujours existé ? En outre Knight néglige le rôle de la détention de capitaux pour expliquer l'existence de l'entrepreneur.

Seconde explication plausible, les agents économiques - producteurs et consommateurs - ont une préférence en soi pour ce mode d'organisation de la production. Coase considère que les hypothèses susceptibles de traduire cette position sont très invraisemblables. Ils imaginent trois types d'hypothèses envisageables : (a) les consommateurs préféreraient les biens produits par les firmes, (b) certains producteurs préféreraient que leur travail soit dirigé par une autorité supérieure, (c) certains auraient un penchant pour le commandement. Coase réfute ces hypothèses selon deux arguments : (1) il existe des firmes dans tous les secteurs, y compris ceux qui produisent des biens pour lesquels les consommateurs ne peuvent qu'être indifférents à la nature de l'organisation de la production, (2) les individus doivent accepter de payer pour satisfaire leur préférence, donc, de rémunérer une autorité supérieure pour qu'elle commande, ou de payer des travailleurs pour qu'ils obéissent. Or, dans la pratique c'est l'inverse que l'on constate : les producteurs préfèrent être leur propre maître et refuseraient de payer pour être dirigés, tandis que ceux qui commandent sont généralement payés pour le faire.

Par conséquent la firme ne peut faire l'objet que d'une préférence relative et non absolue. Dans certaines conditions, la firme s'avère préférable au marché parce que le fonctionnement du marché implique des coûts que la firme entreprise permet de réduire. Ainsi Coase affirme que l'entreprise « *en formant une organisation et en accordant à une autorité le droit de diriger les ressources, certains coûts peuvent être évités.* »⁷. L'échange sur un marché suppose donc des coûts. Ceux-ci sont de trois types :

- **le coût de détermination de prix pertinents.** « *le plus évident des coûts d'organisation de la production par le mécanisme des prix est le coût qu'implique la détermination de prix pertinents.* »⁸ Reste que ce coût n'est pas précisé par Coase. Certes, le mécanisme de détermination des prix est coûteux. Par exemple, pour fonctionner un marché boursier possède des locaux et un réseau informatique dont le coût est répercuté dans chacune des transactions effectuées sur le marché. Autant, pour le marché boursier, il est possible de parvenir à une estimation de ces coûts de détermination des prix, autant pour d'autres marchés, il s'avère extrêmement complexe de déterminer les composants de ce coût de transaction. En outre, la conception du marché n'est pas développée, ce qui aurait pu nous fournir des indications sur ces coûts.

- **les coûts de négociation et de conclusion d'un contrat.** Chaque transaction sur un marché donne lieu à un contrat séparé. La firme passe également des contrats, mais ils sont moins nombreux. « *Un facteur de production n'a pas à passer une série de contrats avec les facteurs avec lesquels il coopère au sein de la firme, comme ce serait nécessaire, évidemment si cette coopération était un résultat*

⁷ *ibid.*, p. 338.

⁸ *ibid.*, p. 336.

direct du fonctionnement du mécanisme des prix. Cette série de contrats est remplacé par un seul. »⁹

En outre, Coase avance un argument complémentaire qui comporte des éléments de coûts, mais renvoie de fait à une autre dimension, l'incertitude (dont il a critiqué l'importance que lui accorde F. Knight dans l'émergence de la firme !).

En effet, Coase avance que, parfois il est préférable, pour la fourniture d'un bien, que les agents établissent un contrat de long terme. Ces contrats évitent aux agents de contracter à chaque transaction. Le choix de ce type de contrat dépend également de l'aversion pour le risque de chacun des agents. L'incertitude est donc une causalité dans l'établissement de relation durable entre agents.

L'incertitude empêche d'établir des prévisions et nécessitent que les agents s'attachent des ressources pour une longue période quitte à définir leur utilisation au cours de son exécution. Comme le note Coase *« Il peut être souhaitable de conclure un contrat à long terme pour la fourniture d'un bien ou d'un service. Ceci peut résulter du fait que certains coûts associés à des contrats individuels peuvent être évités si plusieurs contrats de courte durée sont remplacés par un seul de durée plus longue. OU du fait de l'attitude des parties vis-à-vis du risque, elles peuvent préférer un contrat à long terme plutôt qu'à court terme.... plus le terme du contrat est long, moins il est possible, et en fait moins il est souhaitable, pour l'acheteur de spécifier ce que devra faire l'autre partie contractante. »¹⁰* En revanche, Coase déclare que le vendeur est indifférent au choix que fera ultérieurement l'acheteur dans l'exécution du contrat ; ce qui est pour le moins discutable.

Coase propose donc une première explication à l'existence des firmes dans une économie de marché. Pour être complète, la théorie de Coase doit être en mesure d'expliquer pourquoi certaines transactions s'effectuent uniquement sur des marchés, c'est-à-dire empiriquement, pourquoi la taille des entreprises est limitée ? *« Si par la création d'une firme, on peut éliminer certains coûts et en fait réduire les coûts de production, pourquoi existe-t-il la moindre transaction sur un marché ? Pourquoi la production n'est-elle pas toute entière réalisée par une seule grande firme ? »¹¹*

Pour Coase, la taille de la firme est limitée par le nombre de transactions qu'elle peut intégrer efficacement. Par conséquent, il existera encore des transactions pour lesquelles le mécanisme de marché restera le meilleur mode d'allocation des ressources.

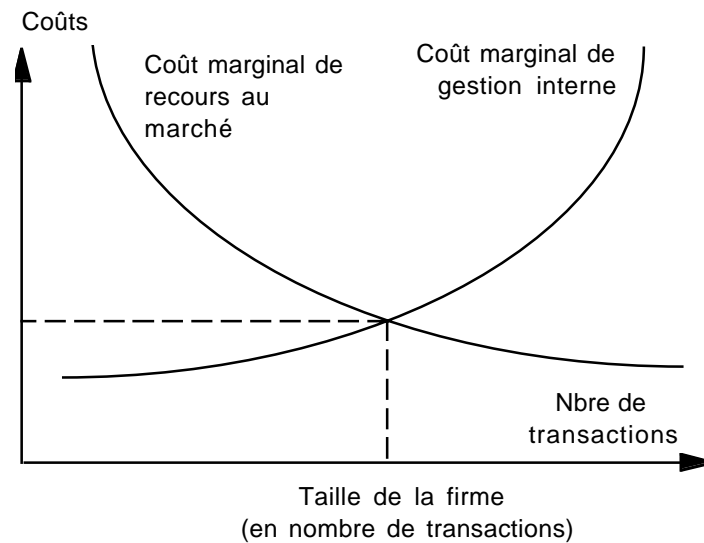
Ainsi, avec la croissance de la firme, *« les rendements de la fonction de l'entrepreneur peuvent être décroissants, c'est-à-dire que les coûts d'organisation de transactions additionnelles au sein de la firme peuvent augmenter. »* Si le coût marginal de gestion interne est croissant, alors il existe naturellement un point ... où les coûts d'organisation d'une transaction supplémentaire à l'intérieur de l'entreprise sont égaux à ceux qu'implique une prise en charge par le marché ou par un autre

⁹ *ibid.*

¹⁰ *ibid.*, p. 337.

¹¹ *ibid.*, p. 340.

entrepreneur. »¹² On peut donc expliquer par cet argument la taille des firmes.(cf. graphique ci-dessous).



Pour Coase, l'augmentation du nombre de transactions gérées en interne peut provoquer des erreurs de la part de l'entrepreneur, ce qui provoque des coûts. « A nouveau, on doit atteindre un seuil où les pertes par gaspillage de ressources sont égales aux coûts de transaction des échanges sur le marché, ou aux pertes subies si un autre entrepreneur organisait la transaction. »¹³ La taille limite de la firme a d'autant plus de chance d'être rapidement atteinte que les coûts de gestion croissent avec la taille des firmes. Ceci est dû à l'augmentation du nombre de transactions mais aussi à l'accroissement de l'hétérogénéité des transactions. Celle-ci entraîne l'augmentation des coûts administratifs du fait de la complexification de la gestion. Par conséquent, il existe une taille maximale à la firme.

L'approche de Coase constitue la base de l'approche des coûts de transaction, mais elle comporte de nombreux problèmes, au premier rang desquels, la définition précise d'un coût de transaction. Coase n'explique pas en quoi consiste réellement les coûts de définition de prix pertinents. Il ne décrit jamais les coûts contractuels et néglige les coûts de contrôle de l'exécution des contrats. Les coûts restent extrêmement vagues dans l'article de Coase. Par ailleurs, le fait qu'aucun mode de gestion des transactions ne s'impose définitivement par rapport à l'autre, suggère que toutes les transactions ne sont pas équivalentes entre elles. O.E Williamson cherchera dans ses travaux à combler ces différentes lacunes et à élargir l'analyse en terme de coûts de transaction. Avant de préciser les apports de Williamson, nous présentons une version formalisée de la relation salariale développée par Herbert Simon. Elle permettra de préciser la notion de hiérarchie introduite par Coase, mais absolument pas développée.

¹² ibid., p. 340.

¹³ ibid., p. 340.

2) La relation d'autorité selon H. Simon¹⁴.

La relation d'autorité est au cœur de la spécificité de l'entreprise en tant que mode d'allocation des ressources. Cette relation a fait l'objet d'une formalisation par H. Simon. Celle-ci sera reprise telle quelle par Williamson dans ses différents travaux. Simon tente de répondre à deux questions :

-1 Dans quelle circonstance est-il mutuellement avantageux pour l'employeur et l'employé de choisir un contrat de travail plutôt qu'un contrat de vente ?

- 2 Pourquoi le salarié accepte-t-il l'autorité de son employeur ?

La réponse de Simon se développe en deux temps. L'incertitude est à la source du choix du contrat de travail. C'est une condition nécessaire mais non suffisante. En cas d'incertitude il est toujours préférable de recourir au contrat de travail ; et l'avantage relatif de ce choix est d'autant plus grand que l'incertitude est forte. L'incertitude n'est pas une condition suffisante dans la mesure où une forme contractuelle ne s'impose uniquement lorsqu'elle maximise la satisfaction **jointe** des deux parties.

Le modèle de Simon.

Soit un employeur B et un employé S. Le salarié S peut *a priori* effectuer un ensemble de tâches X que Simon appelle « ensemble des types de comportements possibles » qui s'identifie de fait à la qualification du salarié considéré. Chaque élément x de cet ensemble est un comportement particulier qui symbolise « *un ensemble de tâches exécutées avec un niveau donné de précisions, d'intensité, etc.* » Simon raisonne à partir de fonctions d'utilité. Les deux parties pour organiser une transaction de prestation de service productif ont deux possibilités :

a) signer un contrat de travail. Cela signifie que S accepte l'autorité de B. Les tâches effectuées (x) par S seront déterminées par B et rémunérées par un salaire (w).

b) signer un contrat de vente par lequel la partie S s'engage à livrer "une quantité déterminée d'une marchandise totalement spécifiée" à l'acheteur B en contrepartie d'une certaine somme d'argent (w). Cet engagement est assimilable à l'acceptation par S d'accomplir un ensemble particulier de tâches (x) pour B.

Les fonctions d'utilité des deux parties, quelque soit le mode d'organisation de la transaction, peuvent donc s'écrire :

$$S_b = F_b(x) - a_b w$$

$$S_s = F_s(x) + a_s w$$

¹⁴ Simon H., [1951], « *A formal theory of the employment relation* », *Econometrica*, Vol. 19, in Simon H., [1957], « *Models of man* », John Willey & Sons, New York.

Pour qu'un accord existe entre les deux parties, il est nécessaire que le contrat de vente ou de travail soit viable pour les deux parties, c'est-à-dire que l'accord sur (x,w) soit avantageux pour les deux parties, soit en terme d'utilité :

$$-F_s(x) \leq a_s w$$

$$F_b(x) \geq a_b w$$

La fonction de satisfaction jointe se définit comme suit :

$$H(x,w) = a_b S_s(x,w) + a_s S_b(x,w) = a_s F_b(x) + a_b F_s(x) = T(x)$$

Simon démontre alors le théorème suivant : l'ensemble des solutions préférées est l'ensemble (x,w) pour lequel $T(x)$ prend sa valeur la plus élevée. Cette valeur est notée T_m . On démontre que :

- (1) si $T(x) = T_m$ pour (x,w) , alors il n' existe pas de point préféré à (x,w) .
- (2) si $T(x') < T_m$ pour (x',w') , alors il existe un point (x,w) avec $T(x)=T_m$ qui est préféré à (x',w') .

Ainsi, le maximum de $T(x)$, satisfaction jointe de B et S, définit un accord optimal, c'est-à-dire un ensemble de tâches et un salaire tel qu'aucun autre accord ne permettrait d'améliorer la situation des deux parties.

Le choix d'un type de contrat dépend de manière cruciale de l'aptitude de B et S à déterminer au moment de la conclusion contrat la satisfaction qui en résultera du choix d'un comportement x pour chacun d'eux. En situation de certitude, c'est-à-dire, lorsque B et S peuvent évaluer avec certitude la satisfaction future qui sera associée à x au moment de la signature du contrat. Dans ce cas là c'est le contrat de vente qui s'impose.

En situation d'incertitude, le choix est beaucoup plus complexe. Les parties devront effectuer des calculs probabilistes et remplacer les résultats certains par les espérances mathématiques . Ou bien les parties connaissent les fonctions de densité de $F_b(x)$ et $F_s(x)$, et ils peuvent déterminer $T(x)$. Dans ce cas, il s'agit d'un contrat de vente en terme d'espérances mathématiques. Ou bien les parties se mettent d'accord sur un salaire et une procédure à suivre ultérieurement, lorsque les valeurs réelles de $F_b(x)$ et $F_s(x)$ seront connues, pour choisir un x particulier. Une des procédures les plus simples consiste pour S à autoriser B à choisir x dans un ensemble X que S accepte. C'est le contrat de travail. Dans ce cas, le choix du contrat de travail est préférable si :

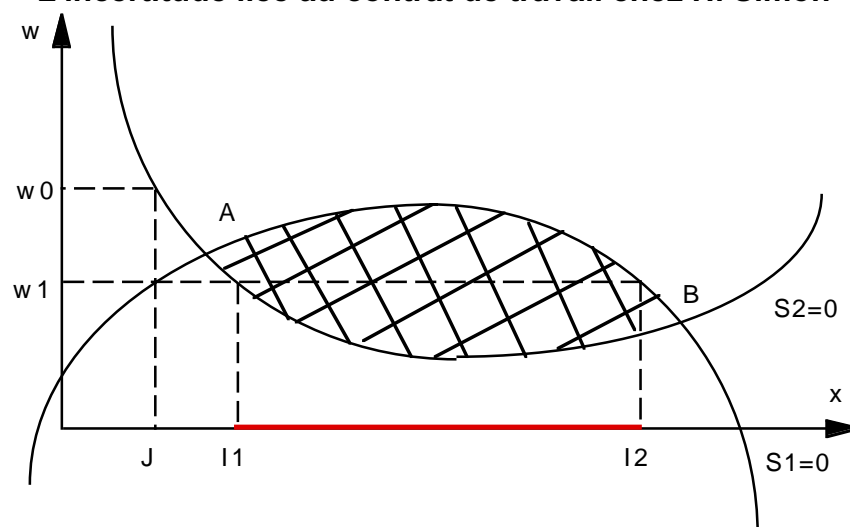
$$a_s F_b \geq a_s a_b w \geq -a_b F_s$$

Cette inégalité ne fait que traduire le fait que l'espace des relations entre le salarié et l'entreprise est naturellement flou, dès lors que l'on raisonne en univers incertain. Entre les parties, il existe une zone acceptable pour les deux parties (*area of acceptance*, zone hachurée sur la graphique page suivante). Le contrat de travail définit ainsi un espace dans lequel l'accord est possible ; il n'y a pas un point d'équilibre comme dans la théorie néo-classique, mais un ensemble de solutions

possibles. Cet espace permet la flexibilité des comportements mais conditionne également l'incertitude. A salaire donné peut correspondre un niveau d'effort plus ou moins important, et donc un ensemble de comportements du salarié acceptables par la firme.

L'intérêt du modèle de Simon est de rendre compte du choix du contrat de travail en situation d'incertitude. Il répond bien aux questions qu'il a posé, sans toutefois précisé suffisamment la seconde. En outre, il néglige totalement le problème de l'exécution du contrat. La question est celle de l'accord sur un type de contrat dont l'exécution est supposée ne pas poser de problème. Or, dans le cas du contrat de travail, il est difficilement envisageable que l'exécution ne pose aucun problème. De fait ce modèle néglige les coûts de fonctionnement d'un système contractuel quel qu'il soit.

L'incertitude liée au contrat de travail chez H. Simon



Pour un effort donné J, le salarié demandera un salaire w_0 et l'entreprise sera prête à verser un salaire w_1 . Il peut y avoir accord en A et B, mais aussi dans toute la zone hachurée. En revanche pour un salaire w_1 , l'effort accepté s'établir entre I1 et I2.

L'apport essentiel de Simon dans l'étude de la relation d'autorité est la prise en compte de l'incertitude dans laquelle évoluent les agents économiques. Cette hypothèse permet de comprendre pourquoi les agents acceptent la relation d'autorité comme mode de coordination des transactions productives. La relation d'autorité sera "déclarée"¹⁵ comme l'organisation du travail la plus efficace par Williamson par rapport à des formes d'organisation individuelle ou collective du travail¹⁶.

Cette hypothèse environnementale couplée à celle de la rationalité limitée constituera le fondement théorique de l'approche de Williamson.

¹⁵ On ne peut à proprement parler de *démonstration* dans le chapitre où Williamson étudie la question, cf. Williamson O.E., [1994], « *Les institutions de l'économie* », pp.249-286, InterEditions, Paris.

¹⁶ op. cit., p. 274.

2) La théorie des coûts de transaction de Williamson.

O.E. Williamson parachève l'œuvre entamée par Coase en lui adjoignant des hypothèses hétérodoxes, au sens où, elles divergent de l'axiomatique néo-classique à laquelle Coase ne déroge pas dans ses travaux. Williamson propose donc une théorie complète des formes d'organisation des transactions qui accepte deux extrema : la firme et le marché. Entre, existent des formes hybrides de gouvernance. La structure de gouvernance qui se mettra en place dépendra des caractéristiques des transactions, principalement de leur fréquence, de la sanction en cas de non respect des engagements. En effet, les agents sont réputés opportunistes et possèdent une rationalité limitée. En outre, les situations de petits nombres sont également étudiées. Nous présentons les différentes hypothèses de Williamson avant de présenter les formes de gouvernance et enfin les formes des firmes.

Les hypothèses du modèle de Williamson : l'homo contractor.

Les hypothèses portent principalement sur les comportements de l'agent économique que Williamson dénomme *Homo contractor*.

La rationalité limitée (*Bounded rationality*)¹⁷.

C'est un concept emprunté à Herbert Simon. Il faut entendre "limitée" non pas dans le sens "irrationnel", mais dans un sens où les individus n'ont pas tous les éléments pour effectuer un choix purement rationnel, c'est-à-dire qui envisage l'ensemble des solutions possibles. Ceci les oblige à se replier sur des solutions qui leur semblent « raisonnable » ou « satisfaisantes ». C'est pourquoi on parle parfois de « satisficing » en lieu et place de « maximizing ». Les individus doivent prendre des décisions dans un contexte flou, ce qui les oblige à opter pour certaines règles ou attitudes qui ne seraient pas les plus appropriées dans une situation où "tout serait parfaitement clair". Plutôt que de rationalité limitée, il vaut mieux parler de **rationalité procédurale**, comme le fait Simon, qui attire l'attention sur le fait que le processus de décision est un élément fondamental de la rationalité.

La rationalité procédurale traduit le fait que les individus ont des capacités cognitives limitées qui les empêchent de calculer *ex ante* l'ensemble des états de la Nature possibles et donc de choisir la meilleure solution. Outre la limitation des capacités cognitives, le traitement de *toute* l'information est impossible. La rationalité telle qu'elle est présente dans le modèle néo-classique est dite substantive, c'est-à-dire que l'on peut procéder à des prédictions à partir des seules données du modèle, le choix des agents ne pouvant être que ceux que le modélisateur peut lui-même établir. A la différence, la rationalité procédurale nécessite de distinguer le modèle du monde réel. On peut distinguer trois types de limites à la rationalité :

¹⁷ Simon H.A., [1972], « *Theories of bounded rationality* », in McGuire C. & Radner R., « *Decision and organization* », American Elsevier, New York ; [1978], « *Rationality as process and as product of thought* », American Economic Review, Vol. 68, pp. 1-16 ; [1986], « *Rationality in psychology and economics* », in The Journal of Business repris dans Hogarth R., Reder M., [1987], « *The contrast between economics and psychology* », Chicago University Press ; [1991], « *Organizations and markets* », Journal of Economic Perspectives.

- le coût de traitement de l'information et la complexité des calculs qui interviennent dans la détermination de la décision "optimale" ;
- les coûts qu'entraîne la collecte de l'information
- l'incertitude inhérente à la vie économique, où le résultat de la décision de chacun dépend de celles des autres.

Simon résume ainsi la démarche qu'implique la prise en compte de la rationalité procédurale : « *Dans les situations qui sont complexes et où l'information est très incomplète (ce qui correspond à pratiquement toutes les situations du monde réel), les approches behavioristes nient qu'il existe une quelconque magie qui engendre des comportements dans lesquels on pourrait voir, même approximativement, une maximisation d'un profit ou d'une utilité. Elles cherchent par conséquent à déterminer le cadre effectif dans lequel la décision est prise, comment ce cadre peut être influencé par elle, et comment elle opère dans le contexte ainsi défini.* »

A cette hypothèse sur les capacités cognitives des individus, Williamson ajoute celle d'opportunisme, hypothèse que l'on déjà rencontrée dans la théorie de l'agence.

L'opportunisme des agents.

L'opportunisme est « *une recherche de l'intérêt personnel qui comporte la notion de tromperie* »¹⁸. Pour Williamson, le comportement opportuniste consiste « *à réaliser des gains individuels dans les transactions par manque de franchise ou d'honnêteté.* »¹⁹ L'opportunisme est inscrit dans la nature humaine et correspond à un comportement stratégique des agents par lequel ils cherchent leur intérêt strictement personnel, quitte à léser l'autre partie au contrat, en recourant à la ruse, la mauvaise foi, le mensonge, la tricherie, etc... Williamson distingue deux types d'opportunisme qui renvoient à l'antisélection et au hasard moral.

- **opportunisme ex ante** : se présente lorsqu'il y a tricherie avant passation du contrat entre les parties. (exemple : communication d'informations erronées par un vendeur). Cet opportunisme est possible à cause de l'asymétrie d'information due à la spécificité des actifs humains (cf. plus loin). Cet opportunisme renvoie au problème d'antisélection (cf. plus haut, Théorie de l'agence, p. 24).

- **opportunisme ex post** : se présente lorsqu'il y a tricherie dans l'exécution du contrat, ou à la fin du contrat (évaluation des actions des parties, question de la reconduction du contrat). Cet opportunisme est lié à l'incomplétude des contrats et à la rationalité limitée, mais aussi à la spécificité des actifs. Cet opportunisme renvoie au problème du hasard moral.

L'hypothèse d'opportunisme ne signifie pas que *tous* les agents sont opportunistes, cela signifie qu'ils *peuvent* tous l'être parce que les individus sont tous censés rechercher leur intérêt personnel. Par conséquent, l'opportunisme introduit un soupçon, un doute, et plus généralement une incertitude comportementale. Cette

¹⁸ Williamson O.E., [1994], *op. cit.*, p. 70.

¹⁹ Williamson O.E., Watcher M., Harris J., [1975], « *Understanding the employment relation : the analysis of idiosyncratic exchange* », Bell Journal of Economics, n°6, p. 258.

incertitude en introduit une autre liée à la stabilité du comportement opportuniste dans le temps : on ne peut savoir à l'avance quel sera le comportement d'autrui. L'introduction du temps - la réalisation des transactions s'effectue dans le temps - rend sensible l'exécution des contrats à l'opportunisme. En effet, si les transactions se dénouaient instantanément, il n'y aurait pas de problème d'opportunisme.

L'introduction du temps et sa prise en considération implicite dans le modèle impliquent que désormais on ne parle plus d'échange mais de transaction. Les transactions peuvent être étudiées en fonction de leurs caractéristiques. Celles-ci sont au nombre de trois.

Les caractéristiques des transactions.

La prise en compte du temps introduit l'incertitude dans les transactions. En outre, la nature des actifs échangés au cours d'une transaction influent fortement sur la forme que va prendre la transaction. Par ailleurs, la fréquence des transactions implique des contrats de nature différente.

• Définition.

La transaction se différencie de l'échange dans la mesure où elle n'est pas instantanée. La transaction est définie de plusieurs façons dans les travaux de Williamson, ce qui ne facilite pas la compréhension exacte de cette notion. Williamson définit la transaction ainsi : « *une transaction a lieu lorsqu'un bien ou un service est transféré à travers une interface technologiquement séparable. Une étape d'activité se termine et une autre commence.* »²⁰. Dans un article antérieur à cette définition, Williamson donnait la définition suivante de la transaction : « *une transaction a lieu lorsqu'un bien ou service est transféré entre deux entités technologiquement séparables* »²¹ Ainsi Williamson donne-t-il le même sens à une relation contractuelle d'achat-vente entre deux entités autonomes juridiquement que le simple transfert entre deux ateliers d'une même usine, transfert qui ne donne lieu à aucun échange monétaire. Les transferts entre entités productives font l'objet de contrats qui définissent « *la tâche particulière qui doit être accomplie* »²² La transaction s'organise en trois temps :

- une négociation initiale;
- un contrat d'échange;
- une résolution du contrat après que l'échange a eu lieu.

La notion de transaction permet de rendre compte de ces moments, en décrivant selon quel processus se réalise l'échange. Ces trois temps sont plus ou moins longs et cruciaux selon la nature des objets échangés. La notion de transaction est donc dans l'approche de Williamson résumée au seul contrat, alors que cette notion avait une toute autre dimension chez les premiers institutionnalistes

²⁰ Williamson O.E., [1985], « *The economic institutions of capitalism* », p. 1, The Free Press, New York.

²¹ Williamson O.E., [1981], « *The modern corporation : origins, evolution, attributes* », p. 1544, Journal of Economic Literature, Vol. 29, Octobre.

²² Williamson O.E., [1985], *op. cit.*, p. 20.

qui avaient introduit cette notion dès les années 1930²³. L'approche des coûts de transaction rapproche ainsi l'économie du droit. « *L'analyse économique des comportements s'enrichit et s'appauvrit en même temps. Elle s'enrichit par ce que, [...], elle intègre les règles de droits des contrats dans le calcul économique et donne de ce fait à l'économie une dimension moins abstraite. Elle s'appauvrit en même temps parce que cette dimension sociale se réduit à l'analyse des règles établissant les contrats, donc à une référence exclusivement juridique de la société.* »²⁴ Comme l'exprime Williamson, « *les coûts de transaction posent le problème de l'organisation économique comme un problème de contractualisation.* »²⁵ La théorie des coûts de transaction exacerbe donc le rôle de l'institution juridico-économique dans le jeu du capitalisme.

L'incertitude des transactions.

L'incertitude est liée à la prise en compte de la rationalité limitée et à l'opportunisme des agents. Pour Williamson, le facteur essentiel d'incertitude repose sur l'opportunisme des agents : « *l'incertitude comportementale, en particulier, a une importance spéciale pour comprendre les sujets étudiés par l'économie des coûts de transaction.* »²⁶ Cette incertitude est qualifiée de stratégique ou comportementale car elle liée à la volonté de certains agents de ne pas divulguer des informations, ou de divulguer une information incomplète et/ou trompeuse. L'incertitude est élevée dès lors que le nombre d'agents est faible. Plus généralement, l'incertitude comportementale n'est pas un problème en soi dès lors qu'il y a suffisamment d'agents substituables entre eux, c'est-à-dire une situation de marché de grand nombre. L'incertitude devient problématique dès lors qu'il y a peu d'agents en mesure de contracter, et c'est très souvent le cas lorsque les actifs sont spécifiques.

La spécificité des actifs.

La nature de l'actif qui fait l'objet d'une transaction a une influence importante sur la forme de la transaction. « *La spécificité des actifs se comprend dans un contexte intertemporel.* »²⁷ En effet, le passé d'un actif, l'expérience acquise avec un actif, limitent sa capacité à être utilisé pour un autre usage. La spécificité des actifs entraîne des inerties et introduit la durabilité des relations. Des actifs dits spécifiques sont difficilement transférables et impliquent des modes de transactions congruents afin d'obtenir une efficacité maximale.

La spécificité d'un actif se définit précisément comme suit : « *des actifs spécialisées [qui] ne peuvent pas être redéployés sans perte de valeur productive en cas d'interruption ou d'achèvement prématuré des contrats.* »²⁸ De fait, la notion d'actif spécifique renvoie à la notion d'actif redéployable. La redepoyabilité d'un actif

²³ cf. Veblen T., [1899], « *Theory of the leisure class* », traduction française 1978, Coll. TEL Gallimard; Commons J., [1934], « *Institutional economics* », Madison, University of Wisconsin Press.

²⁴ Gomez P.Y., [1996], « *Le gouvernement de l'entreprise, Modèles économiques et pratiques de gestion* », p. 69, InterEditions, Paris.

²⁵ Williamson, [1994], *op. cit.*, p. 39.

²⁶ *ibid.*, p. 80.

²⁷ Williamson, [1994], *op. cit.*, p.77.

²⁸ *ibid.*

entraîne des coûts et plus particulièrement des coûts d'opportunité. Les actifs spécifiques entraînent également l'existence de coûts irrécupérables (*sunk costs*)²⁹. La spécificité des actifs renvoie à des investissements durables qui sont réalisés à l'appui de transactions particulières. L'identité des parties importe donc, et la continuité de la relation est valorisée. Par conséquent des garanties contractuelles et organisationnelles particulières sont nécessaires dès lors que les transactions portent sur des actifs très spécifiques. On parle alors de relation idiosyncrasique.

O.E. Williamson distingue quatre types d'actifs spécifiques :

- **la spécificité de site** : la localisation constitue un actif spécifique dans la mesure où elle confère une localisation qui n'est pas reproductible. Ainsi, la localisation d'une entreprise sur une avenue touristique constitue en soi un actif spécifique. Il en va de même si l'on considère deux usines qui pour des raisons de coûts de transport ont installé deux usines côte à côte. Par exemple, la sidérurgie a installé des hauts fourneaux à proximité des mines de charbon. Ce type d'actif spécifique est très peu mentionné par Williamson dans ses travaux à la différence des trois suivants.

- **la spécificité de destination (dedicated assets)** : ce type d'actif spécifique est décrit par Williamson comme la réalisation d'investissement à la demande d'un client unique. Par exemple, l'extension d'une usine existante pour répondre à la demande supplémentaire d'un seul demandeur constitue un actif spécifique. En effet, cet investissement sera difficilement redéployable, c'est-à-dire utilisable à d'autres fins, ou de façon très coûteuse. La mise au point d'une machine spécialisée dans la construction d'une pièce particulière pour un client constitue également un actif spécifique. Cette notion d'actif spécifique permet de comprendre pourquoi les robots industriels sont désormais conçus sur la base de plusieurs opérations, c'est-à-dire que leur usage peut évoluer en fonction des modifications des modèles construits.

- **la spécificité d'actifs physiques** : La mise au point d'une machine spécialisée dans la construction d'une pièce particulière pour un client constitue également un actif spécifique. La différence par rapport aux actifs dédiés est liée à l'existence de caractéristiques particulières du produit qui ne relèvent pas d'une demande particulière mais nécessite un savoir spécifique dans la mise en œuvre.

- **la spécificité d'actifs humains** : elle est liée aux effets d'apprentissage des agents. Ceux-ci deviennent spécialistes dans l'exécution d'une tâche. Cette spécialisation entraîne des difficultés dans la mobilité de ces actifs au sein même des organisations mais aussi entre organisations. Ainsi un mineur de fonds constitue un actif spécifique dans la mesure où il est extrêmement compétent pour travailler dans des gisements de charbon, mais il lui sera difficile - sans coût - de pouvoir effectuer une tâche différente au sein de la même firme, ou au sein d'une organisation alternative.

La fréquence des transactions.

²⁹ les coûts irrécupérables sont des coûts fixes qu'une entreprise ne peut récupérer. Ils sont liés à la spécialité des actifs qui empêchent de fait d'utiliser ces actifs à d'autres fins que celles pour lesquelles ces coûts ont été engagés. Souvent, il n'existe pas de marché d'occasion pour ces actifs.

Cette caractéristique est très importante pour la détermination de la structure de gouvernance des transactions qui va s'imposer aux parties. En effet, le nombre de transactions entraîne des coûts répétés. Si les actifs sont spécifiques, alors de multiples transactions deviennent très coûteuses. Il devient judicieux de les regrouper au sein d'une organisation afin d'économiser des coûts de transaction. En outre, si les transactions sont peu nombreuses, cela signifie que le nombre d'acteurs est faible. Par conséquent, l'opportunisme a plus de chance de se développer et augmente l'incertitude qui pèse sur la transaction.

Les coûts de transaction.

Williamson précise un peu plus la notion de coûts de transaction que Ronald Coase. Il distingue ainsi les coûts de transaction *ex ante* des coûts de transactions *ex post*.

Les coûts de transaction *ex ante* comportent :

- les coûts de négociation du contrat ;
- les coûts de rédaction du contrat ;
- les coûts de garanties.

Ces dernières peuvent être très complexes et Williamson insiste sur ce type de coûts. En effet, la garantie habituellement avancée par les économistes est le recours au tribunal. Williamson avance que ce recours peut être coûteux sinon inefficace. Par conséquent, les parties s'entendent sur de mécanismes d'arbitrage (parfois complexes) dont le rôle est de se substituer au mécanisme que constitue la voie juridique classique.

Les coûts de transaction *ex post* sont constitués par :

- les coûts de mauvaise adaptation du contrat ;
- les coûts de marchandage occasionnés par la divergence dans l'exécution du contrat ;
- les coûts d'organisation et de fonctionnement associés aux structures de gouvernance auxquelles les parties s'adressent en cas de conflit ;
- les coûts d'engagements sûrs (*credible commitments*).

Les coûts *ex ante* et *ex post* sont interdépendants et doivent être pris en considération de façon globale et non séquentielle. Ils sont extrêmement complexes à quantifier. Mais Williamson avance que leur évaluation en tant que telle importe peu puisque dans un optique comparative des structures de gouvernance, seule la différence de coûts entre deux structures de coût importe. Comme l'écrit Williamson : « *Les recherches empiriques sur les coûts de transaction n'essaient presque jamais de mesurer directement de tels coûts. Au lieu de cela, la question est de savoir si les relations organisationnelles (pratiques contractuelles, structure de gouvernance) s'ajustent ou non avec les attributs des transactions que prédit le raisonnement des coûts de transaction.* »³⁰

³⁰ Williamson O.E., [1994], *op. cit.*, pp. 41-42.

L'objectif est de donc de parvenir à trouver une structure de pilotage des transactions qui minimisent à la fois les coûts de transactions et les coûts de production de l'actif sous-jacent à la transaction : « *Il faut donc évaluer si les économies de coûts de transactions sont réalisées aux dépens des économies d'échelle ou de gamme. Un cadre d'arbitrage est nécessaire pour examiner simultanément les implications en terme de coûts de production et de coûts de gouvernance des formes alternatives d'organisation.* »³¹

Les structures de gouvernance efficaces.

Les caractéristiques des transactions vont directement déterminer les structures de gouvernance efficaces, c'est-à-dire celles qui minimisent à la fois les coûts de transaction et les coûts de production. Trois caractéristiques déterminent les structures de gouvernance : la spécificité des actifs, l'incertitude et la fréquence des transactions.

Avant de présenter les structures de gouvernance des transactions efficaces nous présentons différents exemples de transactions qui permettent d'éclairer la grille d'analyse des coûts de transactions.

Caractéristique des actifs

Fréquence des transactions (ci dessous)	Non spécifique	Mixte	Idiosyncrasique
Occasionnelle	Achat d'équipements standards	Achat d'équipements sur commande	Construction d'une usine
Récurrente	Achat de matériaux standards	Achat de matériaux sur commande	Transfert spécifique en site de produits intermédiaires à travers des étapes successives

Source : Williamson , [1994].

Le croisement des trois caractéristiques permet de mettre en lumière des situations contractuelles différentes pour lesquelles il existe une structure de gouvernance adéquate. Les types de structures de gouvernance peuvent être regroupés en trois modes de contractualisation : classique, néo-classique, évolutive.

³¹ *ibid.*, p. 83.

Les structures de gouvernance efficaces (Williamson 1985)

		DEGRE DE SPECIFICITE DES ACTIFS		
		Non spécifique	Mixte	Idiosyncrasique
FREQUENCE DES TRANSACTIONS	OCCASIONNELLE	GOUVERNANCE DU MARCHÉ <i>(Contractualisation classique)</i>	GOUVERNANCE TRILATERALE <i>(Contractualisation néo-classique)</i>	
	RECURRENTE		GOUVERNANCE BILATERALE <i>(Contractualisation évolutive)</i>	GOUVERNANCE UNIFIEE

La gouvernance de marché (contractualisation classique). Ce mode de pilotage des transactions correspond à l'échange tel que le représente l'analyse économique néo-classique. Les transactions sont ponctuelles et l'objet des transactions est parfaitement délimités, c'est-à-dire que ces caractéristiques sont connues et objectivables. Il n'y a pas d'incertitude : toutes les éventualités sont prévisibles. Il n'y a donc pas de problème de contractualisation ; le contrat envisage toutes les éventualités qui pourraient survenir. L'identité des parties importe peu dans la réalisation de la transaction. L'opportunisme des agents a peu de chance de se développer dans la mesure où, la non spécificité des objets échangés implique qu'il existe un grand nombre d'offreurs possibles. Le grand nombre d'offreurs limite l'opportunisme car un client insatisfait ou trompé par un offreur opportuniste pourra le sanctionner en n'effectuant son achat suivant auprès d'un autre offreur, ce qui ne serait pas possible dans une situation de petit nombre d'agents.

La gouvernance trilatérale (contractualisation néo-classique). Dès que la spécificité des actifs s'élèvent, la durée de la relation importe du fait de l'existence de coûts irrécouvrables générés par la spécificité des actifs. En cas de rupture du contrat, les parties subissent une perte importante. Par conséquent les parties ont intérêt à ce que la relation soit durable. Les parties s'engagent donc mutuellement au travers de contrats de long terme. Mais du fait de la rationalité limitée des individus et de l'incertitude comportementale liée à l'opportunisme des agents, le contrat de long terme ne peut *ex ante* prévoir toutes les occurrences qui vont influencer sur la réalisation du contrat. Afin de gérer les conflits en cours d'exécution du contrat, le contrat inclut une instance d'arbitrage qui n'est pas une instance juridique. En effet, pour Williamson, le recours à la Justice n'est pas forcément efficace, dans la mesure où il

est long et coûteux. Qui plus est, la spécificité accrue des actifs sous-jacents à la transaction nécessite souvent des compétences d'expertise que les tribunaux ne possèdent pas forcément. C'est pourquoi la structure de gouvernance de ce type de transaction est un contrat qui inclut une instance d'arbitrage ; d'où le nom de gouvernance trilatérale. La faible occurrence des transactions explique pourquoi les contractants ne recourent pas à un mécanisme de pilotage durable (firme par exemple). Ce type de solution serait trop coûteuse compte tenu de la faible occurrence de la transaction.

La contractualisation évolutive. Elle s'impose lorsque la spécificité des actifs est très élevée et que la complexité des liens entre les parties qu'elle engendre, devient importante. En effet, la forte spécificité des actifs implique une étroite collaboration entre les parties. Ici, l'identité des parties importe au plus haut point. Il se noue ainsi des liens durables entre les contractants. Le choix entre la structure bilatérale et la structure unifiée - i.e. la firme - dépend principalement de la spécificité des actifs. La différence entre les deux structures repose sur l'autonomie des contractants dans le contrat.

- **Gouvernance bilatérale** : l'autonomie des parties est maintenue. On peut regrouper dans ce type de relations contractuelles les accords de partenariat entre firmes, les contrats de sous-traitance, les alliances entre firmes. Le choix entre cette structure et l'intégration verticale s'effectuera sur la base de l'adaptabilité de ces structures aux aléas qui pèsent sur la transaction et sur la base de la récupération des économies d'échelle que peut engendrer le recours à la sous-traitance.

- **Gouvernance unifiée** : la transaction est retirée du marché et organisée au sein de la firme pour être soumise à une relation d'autorité. C'est l'intégration verticale. La présence d'actifs très spécifiques (humains et physiques) permet l'existence d'économies d'échelle. En effet, la spécialisation des actifs et l'apprentissage de leur usage permet de dégager des économies d'échelle. Par conséquent la fréquence élevée des transactions et la forte spécificité des actifs expliquent l'existence de la **firme**.

Le même outil d'analyse - le coût de transaction - permet donc d'expliquer à la fois l'existence du marché et de la firme, mais aussi de formes hybrides (structure de gouvernance entre marché et firme).

L'intégration verticale.

O.E. Williamson précise un peu plus la question de l'intégration verticale.

Soit dG la différence de coût entre les coûts de la relation d'autorité et les coûts de recours au marché :

$\Delta G = \text{Coût de la bureaucratie} - \text{Coût de recours au marché.}$

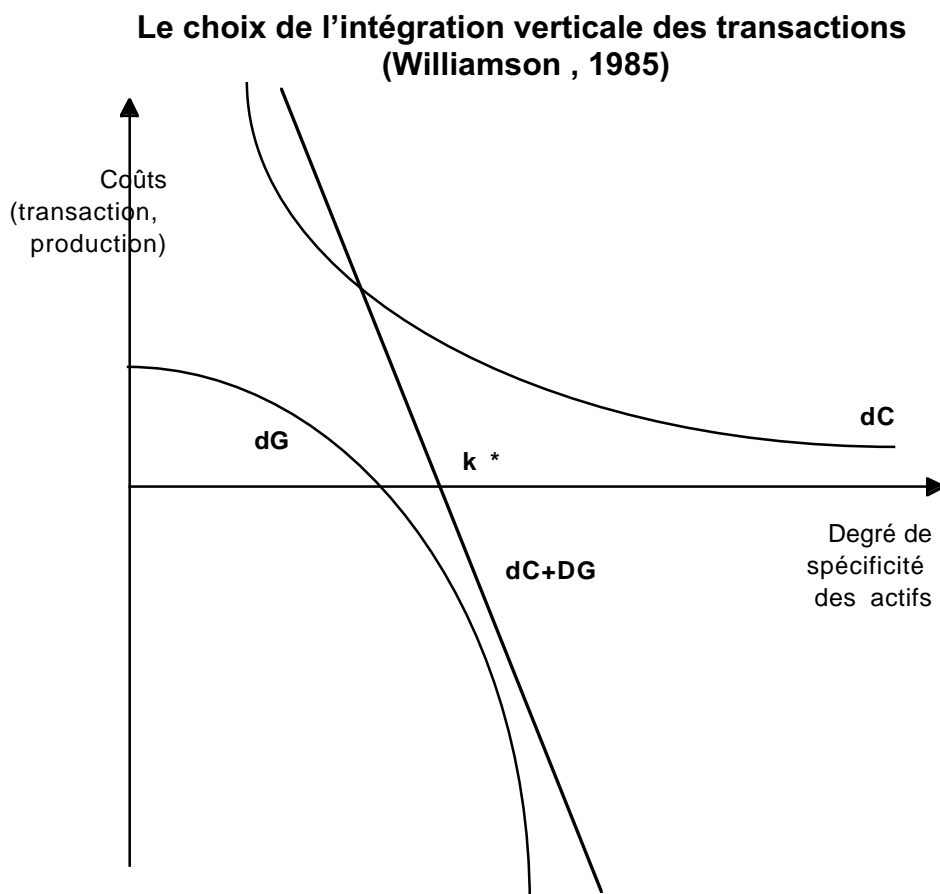
Si $\Delta G < 0$, alors il est préférable d'intégrer la transaction au sein de la firme, puisque le coût interne est inférieure à celui du marché. Cette fonction est décroissante du degré de spécificité des actifs. En effet, plus les actifs sont spécifiques, plus la firme est adaptée à la gestion des transactions, et plus les avantages à l'externalisation diminuent. Il faut donc internaliser la transaction. Mais

cette première analyse n'est pas suffisante dans la mesure où il faut minimiser à la fois les coûts de transactions et les coûts de production classiques.

Soi dC la différence de coût de production entre la firme et la production sur le marché. Ce différentiel ΔC est strictement positif car le recours au marché peut permettre de bénéficier d'économies d'échelle. Ce différentiel est une fonction décroissante du degré de spécificité des actifs pour les mêmes raisons que précédemment.

C'est donc la somme de ces deux différentiels qu'il faut prendre en considération pour savoir si la transaction doit être retirée du marché ou non. Si $\Delta G + \Delta C < 0$ alors l'intégration verticale devient préférable. Il existe un point k^* où on est indifférent entre la solution de marché et la solution d'intégration (cf. le graphique page suivante).

Pour Williamson ce sont donc les deux sortes de coûts qu'il faut prendre en considération dans le choix d'internalisation / externalisation de la transaction. Ceci constitue un renouvellement important de l'analyse du phénomène de l'intégration verticale. Mais, il est paradoxal de constater que Williamson ne parle quasiment que des économies d'échelle dans cette discussion, pour finalement conclure laconiquement que ce sont les coûts de transaction qui sont les plus déterminants de le choix de la structure de gouvernance !



Le concept de coût de transaction constitue donc apparemment un outil d'analyse puissant pour expliquer l'existence de structures différentes de

coordination des transactions. Cet outil d'analyse permet même de mettre en lumière l'efficacité des formes de gouvernance unifiée les plus efficaces. Ainsi, Williamson est-il en mesure de proposer une structure organisationnelle efficace pour gérer les transactions en interne.

Ainsi, la théorie des coûts de transaction est-elle en mesure de proposer une grille d'analyse qui permet à la fois d'expliquer l'existence de modes de coordination différents en fonction de deux critères principaux : la spécificité des actifs et la fréquence des transactions.

La firme est assimilée à un nœud de contrats. La frontière de la firme est déterminée par la prise en compte des coûts de transaction et des coûts de production. Cette théorie propose ainsi un continuum entre la firme et le marché en introduisant des structures de coordination hybrides (gouvernance bilatérale). Mais dans le même temps, en résumant la firme à un nœud de contrats, elle détruit toute spécificité de la firme par rapport au marché, et de fait, parler de frontière de la firme n'a pas grand sens dans cette théorie puisque toute transaction est contrat, tant au sein de la firme, tant sur un marché.