

COLLECTION « ÉCONOMIE ET STATISTIQUES AVANCÉES »

Jean TIROLE

**THÉORIE
DE L'ORGANISATION
INDUSTRIELLE**

TOME I

Paris, Economica Éditions, 2013.

contraste, un employé embauché pour une tâche particulière, et un autre pouvant démissionner ou être renvoyé sans préavis et sans indemnités de licenciement).

La troisième vision de l'entreprise - celle des contrats incomplets - s'approche un peu plus de la (ou des) définition(s) juridique(s). Elle souligne le fait que les entreprises et les contrats sont des « modes de gouvernement » assez différents. Elle considère l'entreprise comme une façon particulière de spécifier ce qui doit être fait si des événements contingents surviennent, qui ne sont pas prévus dans le contrat. L'on part de l'idée que les contrats sont nécessairement incomplets, parce que certains événements contingents sont imprévisibles, ou parce qu'il en existe trop pour les spécifier tous par écrit, de sorte que la minimisation du coût exige que le contrat original ne définitive que les grandes lignes de la relation entre les parties. La propriété des machines et la relation d'emploi donnent au propriétaire de l'entreprise le pouvoir (« autorité ») de choisir la façon d'employer les machines et le personnel, sous certaines limites. Le degré d'intégration peut alors être grossièrement mesuré par le degré selon lequel l'autorité est répartie entre les contractants.

Sont également traités dans la section 1 les facteurs limitant la taille de l'entreprise, en fonction de ses différentes définitions. Parmi ces facteurs figurent l'épuisement des économies d'échelle et de taille, les facteurs rares utilisés, les inconvénients des relations de long terme et la concentration de pouvoir excessive.

La section 2 considère les entreprises comme des entités bien définies. Plutôt que d'analyser leur taille, elle s'interroge sur la façon dont les entreprises se comportent. Les entreprises maximisent-elles vraiment leurs profits, comme supposé dans tout ce livre ? Si tel n'est pas le cas, l'hypothèse de maximisation du profit est-elle encore un guide utile pour la détection d'un comportement monopolistique ou d'une interaction stratégique ? Les différents mécanismes qui mettent un frein au pouvoir discrétionnaire des dirigeants d'entreprise (le contrôle ; les incitations sur les marchés du produit, du travail et du capital) sont passés en revue, et le postulat de maximisation du profit est alors discuté.

1 Qu'est-ce qu'une entreprise ?

1.1 L'entreprise comme une échappatoire à l'interdiction du pouvoir de monopole

Plusieurs pratiques commerciales sont, justement ou injustement, proscrites par la réglementation antitrust, car considérées comme des manifestations perverses d'un pouvoir de monopole. En internalisant de telles pratiques, les entreprises peuvent contourner le cadre juridique et exercer tranquillement leur pouvoir de monopole. (Les transactions internes étant habituellement non observées, elles échappent à la loi¹).

La discrimination par les prix

Comme on le verra dans le chapitre 3, une entreprise peut vouloir vendre le même produit sur différents marchés à des prix différents. Cela crée la possibilité d'un arbitrage par les grossistes fournissant ces marchés. Pour éviter cet arbitrage, le fabricant peut intégrer le secteur de la distribution et fournir lui-même le marché concerné par le bas prix. Un phénomène semblable, induit par la loi plutôt que par l'arbitrage, se produit lorsque la justice contraint une entreprise à traiter deux clients dans une « situation similaire » de façon symétrique (ce qui est en principe le cas aux Etats-Unis ou en France), alors que le fabricant souhaite les traiter différemment, ou restreindre l'accès de l'un des clients. En intégrant un des clients, le fabricant peut traiter l'autre différemment, ou bien encore probablement fermer tout simplement le marché du bien intermédiaire.

Les contrôles des prix intermédiaires

Supposons que le prix d'un bien intermédiaire, échangé entre un grand nombre d'offres et un grand nombre de demandeurs, soit fixé par des agences de régulation à un niveau inférieur au prix d'équilibre du marché, déterminé par l'offre et la demande. Le bien doit alors être distribué aux acheteurs par un mécanisme de rationnement. Cela donne aux offres une incitation à fusionner avec certains de leurs acheteurs. Pourquoi ? Parce que le prix de la transaction (légal) portant sur le bien intermédiaire est inférieur au « prix implicite » que les acheteurs accepteraient de payer. Par conséquent, il existe des rentes potentielles, que pourraient se partager les offres et les utilisateurs rationnés. Cependant, les acheteurs qui ne peuvent s'approprier leur part de rentes en augmen-

tant leur prix quand le marché est libre, ne peuvent plus le faire quand celui-ci est réglementé. L'intégration verticale permet alors aux offreurs de contourner la loi en engendrant des transactions internes - et donc non observables.

Il y a d'autres raisons juridiques qui peuvent conduire des entreprises liées verticalement à s'intégrer. L'une d'entre elles est l'existence d'un impôt sur les ventes. L'internalisation des transactions évite le paiement d'une telle taxe au niveau intermédiaire. (Ce phénomène ne se produit pas lorsque l'impôt est calculé sur la valeur ajoutée plutôt que sur la valeur totale du produit au moment de la transaction. Un impôt sur la valeur ajoutée est neutre par rapport à la décision d'intégration). Une autre raison est l'existence d'une régulation par le taux de rendement. Une entreprise sujette à une telle régulation peut augmenter ses profits en intégrant en amont dans le secteur des biens d'équipement, si les prix de ce secteur ne sont pas réglementés¹.

Le chapitre 4 est consacré à l'étude des motivations économiques plutôt que juridiques de l'intégration verticale, reliées à l'exercice du pouvoir du monopole.

Remarque : L'intégration horizontale peut également être motivée par la volonté d'exercer un pouvoir de monopole. Un exemple évident est celui d'un regroupement de deux entreprises produisant les mêmes biens. De tels regroupements, qui ont dominé durant la vague de fusions de 1887-1904 aux États-Unis², sont souvent destinés à récupérer les profits de monopole qui sont dissipés du fait de la concurrence sur le produit (voir le chapitre 5)³.

1.2 L'entreprise en tant que synergie statique

Un thème classique de l'organisation industrielle, dérivé de l'étude bien connue de Viner (1932) sur les courbes de coût, est que la taille et le nombre des entreprises dans un secteur d'activité sont liés à l'ampleur des économies d'échelle.

L'un des principaux déterminants de la taille d'une entreprise est le degré avec lequel elle peut exploiter les économies d'échelle

1. L'entreprise réglementée peut gonfler artificiellement les prix de transfert afin de relâcher la réglementation par le taux de rendement. Voir Dayan (1972).

2. Voir par exemple Scherer (1980) : pp. 119-122. Des fusions de ce type ont ensuite été découragées par l'application de lois antitrust.

et de taille. Comme cela a bien été montré par les ingénieurs¹, des niveaux de production plus élevés permettent l'utilisation de techniques plus efficaces. Ils justifient d'investir dans des technologies de réduction des coûts et permettent aux travailleurs d'être plus spécialisés. Les coûts unitaires diminuent. De telles économies d'échelle, liées au volume d'activité sur un bien unique, sont appelées économies spécifiques au produit. Bien que cela soit inhabituel on peut inclure dans cette catégorie ce qu'on appelle les économies de réserves groupées (Robinson, 1958). Une usine ayant un grand nombre de machines peut soutenir un flux de production proportionnellement plus important qu'une usine n'ayant qu'un petit nombre de machines. Ceci est dû au fait que la panne aléatoire d'une machine a moins d'impact sur la production, puisque le flux de production destiné à cette machine peut être dirigé vers les autres machines. (A la limite, avec beaucoup de machines, et si les pannes de ces machines sont indépendantes, le flux de production soutenu est parfaitement déterministe, en raison de la loi des grands nombres). De façon analogue, une entreprise vendant sur plusieurs marchés où les demandes sont variables (et parfaitement corrélées) rencontre moins d'incertitude qu'une multitude de firmes séparées servant ces mêmes marchés indépendamment, et donc peut économiser de coûteux investissements de capacité de pointe. Les économies de réserves groupées peuvent aussi s'appliquer aux entreprises multiproduits, lorsque les différents produits partagent les mêmes techniques de production².

D'un point de vue différent, le groupement d'activités, qu'elles dépendent du même produit ou de produits différents, peut être lié non pas à la production au sens étroit, mais à tous les services qui l'accompagnent : audit, marketing, gestion du personnel, finance, distribution, fourniture de matériaux, recherche et développement, et ainsi de suite. Un tel regroupement évite le double emploi des coûts fixes associés à ces services, ou du moins réduit ces coûts en moyenne. Les complémentarités de demande peuvent également être un motif de coordination d'activités³.

1. Voir Scherer (1980), pp. 81-84 pour quelques exemples concrets.

2. Un article utile pour approfondir les raisonnements de ce paragraphe est celui de Arrow *et al.* (1972). Les auteurs montrent le lien entre la réparation des machines et les rendements d'échelle croissants. Ces auteurs obtiennent une fonction de production à rendements d'échelle croissants, qui tend vers une fonction de production à rendements d'échelle constants, lorsque le nombre de

Annexe :

Une démonstration heuristique de la conjecture de Coase

La démonstration heuristique suivante, qui est destinée au lecteur avancé, est inspirée de Wilson (1985).

Considérons un horizon infini $t = 1, 2, \dots$. On peut facilement montrer que pour toute suite de prix passés fixés par le monopoleur, $\{p_1, \dots, p_{t-1}\}$, les croyances *a posteriori* du vendeur au début de la période t sont définies ainsi : il existe un nombre tel que les acheteurs avec des évaluations b appartenant à $[0, b_t]$ n'ont pas encore acheté le bien, alors que les acheteurs avec des évaluations b appartenant à $]b_t, +\infty[$ l'ont déjà acheté. En d'autres termes, les croyances *a posteriori* du monopoleur sont nécessairement obtenues par troncature de ses croyances *a priori* (ceci provient du fait qu'un acheteur avec une évaluation élevée est plus impatient d'acheter). En suivant Gul et al. (1986), supposons que les acheteurs suivent une stratégie « stationnaire » simple. Quand on leur fait payer p_t à la date t , les acheteurs avec des évaluations dépassant $\beta(p_t)$ achètent (s'ils ne l'ont déjà fait), alors que ceux qui ont une évaluation inférieure à $\beta(p_t)$ n'achètent pas, où $\beta(\cdot)$ est une fonction croissante telle que $\beta(p_t) > p_t$ pour tout $p_t > 0$. (L'importance de cette hypothèse de stationnarité a été démontrée par Ausubel et Deneckere, 1986). Pour simplifier les notations, posons $c = 0$ le coût marginal du monopoleur et $F(b)$ la fonction de répartition des évaluations des acheteurs sur $[0, +\infty[$ (avec $F(0) = 0$, $F(b) > 0$ pour $b > 0$ et $F(+\infty) = 1$). Enfin, $\delta = \exp(-r\Delta)$ représente le facteur d'escompte, où Δ est la longueur de temps réel entre les périodes. Nous nous intéressons à ce qui se passe lorsque Δ tend vers 0.

En raison de la stationnarité de la stratégie des acheteurs, la valeur actualisée courante des profits du monopoleur à partir d'une certaine date t ne dépend que de la distribution des acheteurs restants, ce qui est résumé par « l'évaluation critique » b_t . Soit $V(b_t)$ ce profit actualisé courant. Remarquons que $V(\cdot)$ est une fonction non décroissante, et posons $F_t \equiv F(b_t)$, la proportion d'acheteurs qui n'ont pas acheté avant la date t .

Fixons un temps réel $\varepsilon > 0$, et laissons Δ converger vers 0. Pour un $\eta > 0$ quelconque, il existe un Δ suffisamment petit et un t (tel que $(t+2)\Delta < \varepsilon$) satisfaisant

$$F_t - F_{t+2} \equiv F(b_t) - F(b_{t+1}) < \eta \quad (A1)$$

En d'autres termes, lorsque le nombre de périodes entre 0 et t tend vers l'infini, on peut toujours trouver deux périodes consécutives telles que la quantité totale vendue durant ces deux périodes est bornée par un nombre donné.

L'intuition de Coase était que si le profit $V(b_{\varepsilon/\Delta})$ à partir du temps réel ε était non négligeable, le monopoleur serait incité à accélérer le processus en réduisant plus vite le prix. Par exemple, entre les dates t et $t+2$ il ne vend pas beaucoup et ne profite pas beaucoup de la discrimination non plus, car le temps réel entre ces deux périodes est court (de sorte que le prix ne peut décroître beaucoup entre ces périodes ; sinon les acheteurs attendraient et aucun n'achèterait entre ces deux périodes). Par conséquent, en un sens, le gain issu de la discrimination est de second ordre, alors qu'en faisant à la date t l'offre qu'il projette de faire à la date $t+1$, le monopoleur accélérerait le processus d'une période et obtiendrait un gain de premier ordre (en Δ) si $V(b_{\varepsilon/\Delta})$ n'est pas trop petit. Pour formaliser cette intuition, écrivons la condition selon laquelle le monopoleur préfère faire payer p_t à la date t et p_{t+1} à la date $t+1$ plutôt que p_{t+1} directement à la date t :

$$p_t (F_t - F_{t+1}) + \delta p_{t+1} (F_{t+1} - F_{t+2}) + \delta^2 V(b_{t+2}) \geq p_{t+1} (F_t - F_{t+2}) + \delta V(b_{t+2}) \quad (A2)$$

cela est équivalent à

$$(p_t - p_{t+1})F_t - (p_t - \delta p_{t+1})F_{t+1} + (1 - \delta)p_{t+1}F_{t+2} \geq \delta(1 - \delta)V(b_{t+2}) \quad (A3)$$

Par définition de b_{t+1} , l'acheteur avec l'évaluation b_{t+1} est indifférent entre accepter p_t et accepter p_{t+1} ; par conséquent,

$$b_{t+1} - p_t = \delta (b_{t+1} - p_{t+1}) \quad (A4)$$

ce qui implique que

$$p_t - \delta p_{t+1} = (1 - \delta)b_{t+1} \quad (A5)$$

et

$$p_t - p_{t+1} = (1 - \delta)(b_{t+1} - p_{t+1}) \quad (A6)$$

En substituant les équations A5 et A6 dans l'équation A2 et en divisant par $(1 - \delta)$, nous obtenons

$$(b_{t+1} - p_{t+1})F_t - b_{t+1}F_{t+1} + p_{t+1}F_{t+2} \geq \delta V(b_{t+2}) \quad (A7)$$

c'est-à-dire

$$b_{t+1}(F_t - F_{t+1}) - p_{t+1}(F_t - F_{t+2}) \geq \delta V(b_{t+2}) \quad (A8)$$

Mais $t+2 \leq \varepsilon/\Delta$ implique que $b_{t+2} \geq b_{\varepsilon/\Delta}$ et donc que $V(b_{t+2}) \geq V(b_{\varepsilon/\Delta})$. On a aussi $F_{t+1} \geq F_{t+2}$. Par conséquent, l'équation A8 devient

$$(b_{t+1} - p_{t+1})(F_t - F_{t+2}) \geq \delta V(b_{\varepsilon/\Delta}) \quad (A9)$$