Essais en Théorie des Organisations : Incitations et Structure des Organisations

Radoslava Nikolova

Université de Montpellier I
Faculté des Sciences Économiques

Année 2007

Numéro attribué par la bibliothèque
| | | | | | | | | | | | |

THÈSE

Pour l'obtention du grade de
Docteur de l'Université de Montpellier I
Discipline : Sciences Économiques

Présentée et soutenue publiquement par

Radoslava Nikolova

le 14 Décembre 2007

Essais en Théorie des Organisations :
Incitations et Structure des Organisations

Directeur de thèse : Wilfried Sand-Zantman

JURY :

Monsieur Bernard Caillaud (rapporteur)	Ingénieur en Chef des Ponts et Chaussées
Monsieur Wouter Dessein	Professeur à l'Université de Chicago
Monsieur Robert Gary-Bobo (rapporteur)	Professeur à l'Université de Paris I
Monsieur Bernard Salanié	Professeur à l'Université de Columbia (New York)
Monsieur Wilfried Sand-Zantman	Professeur à l'Université de Toulouse I
Monsieur Marc Willinger	Professeur à l'Université de Montpellier I

L'université Montpellier I n'entend donner aucune approbation ni improbation aux opinions émises dans cette thèse. Ces opinions doivent être considérées comme propres à leur auteur.

Remerciements

Je remercie mon directeur de thèse Wilfried Sand-Zantman pour m'avoir donné envie de faire de la recherche, pour avoir accepté de diriger cette thèse, pour ses conseils, ses encouragements et son soutien tout au long de ces quatre années.

Je suis très reconnaissante envers Bernard Salanié pour sa très grande disponibilité et sa patience à mon égard, pour les nombreuses discussions et toutes les choses qu'il m'a apprises, pour son exigence et sa rigueur qui m'ont faite constamment avancer.

Je souhaite également remercier tous les membres de mon jury, Bernard Caillaud, Wouter Dessein, Robert Gary-Bobo, et Marc Willinger d'avoir accepté d'en faire partie.

Tout au long de cette thèse j'ai bénéficié d'un cadre de travail exceptionnel au sein du Laboratoire d'Économie Industrielle. Mes remerciements s'adressent en particulier à Anne Perrot pour m'y avoir accueillie, à Jérôme Pouyet pour son écoute et ses conseils lors de mes débuts de thèse hésitants (et après), à Philippe Fevrier pour sa rigueur et son dynamisme, dont j'ai pu bénéficier au cours des derniers mois. Merci aussi à Jacquie pour son aide lors de diverses démarches administratives et pour son sourire très chaleureux. Ma dernière année de thèse a été financée grâce à l'allocation d'ATER de l'Université Cergy-Pontoise. Je remercie en particulier mon collègue d'enseignement David Ettinger, pour son accueil à Cergy et pour m'avoir proposé de partager ses enseignements.

Au cours des dernières semaines plusieurs personnes ont accepté de relire des morceaux de la thèse. Je remercie Basak Bayramoglu, Thomas Tregouet, Laurent Lamy, Thibaud Vergé, Jean-François Jacques, Sylvaine Poret, Philippe Fevrier, Pierre Cahuc, Laure Durand-Viel, Jeanine Miklos-Thal, Laurent Linnemer et Arnold Vialfont pour leurs conseils et critiques.

La thèse a été l'occasion de nombreuses rencontres. Je remercie en particulier Sylvaine, Jorge, Hélène, Johan, Lionel, Thibaud, Thomas C., Laure, Arnold et Gautier pour avoir égayé les déjeuners et pauses café.

Thomas et Laurent, merci pour votre bonne humeur et pour toutes les discussions, certaines sérieuses d'autres moins, j'ai sincèrement apprécié. Un grand merci à Anne, Basak et Jeanine, pour les nombreux bons moments passés ensemble et pour leur capacité à toujours trouver les mots que j'ai besoin d'entendre. Mention spéciale pour Basak qui m'a supportée au quotidien pendant cette dernière année et a géré mes coups de stress. Je remercie également Benoit de toujours arriver à me faire voir la vie en rose. Je pense aussi à tous mes amis qui ont été là de près ou de loin et qui ont vite su abandonner la question "Alors, ça avance ?".

Je remercie mes parents et mon frère pour leur confiance de toujours et leur soutien inestimable. Merci aussi à Catherine et Michel. Enfin, merci à Philippe d'être à mes côtés.

Sommaire

Introduction générale

> « *The economies of modern industrialized society can more appropriately be labeled organizational economies than market economies. Thus, even market driven capitalist economies need a theory of organizations as much as they need a theory of markets.* »
>
> Herbert Simon, *Journal of Economic Perspectives*, 1991, p. 42

Dans les années 1920, sous la direction d'Alfred Sloan, General Motors[1] entame un changement organisationnel majeur. L'entreprise se tourne vers une structure multidivisionnelle[2]. Cette nouvelle forme organisationnelle va la propulser, au début des années 1930, au premier rang des entreprises industrielles. En 1940 sa part de marché est de 45% (contre 11% en 1921), loin devant l'ancien leader de l'industrie Ford (16%). Ainsi, l'adoption de la forme multidivisionnelle est qualifiée d'invention majeure du vingtième siècle et laisse peu de doutes quant à l'importance de la forme organisationnelle sur la performance de l'entreprise. Cependant, l'intérêt des économistes pour les limites de l'entreprise, son organisation, ses systèmes d'information et d'incitation est relativement récent. Pendant longtemps l'analyse de l'entreprise, dans la théorie microéconomique, était réduite à sa fonction de production. Coase (1937) est à l'origine d'une lente mais néanmoins importante

[1] Parmi d'autres grandes entreprises comme Du Pont de Nemours, Standard Oil et Sears. Pour une analyse détaillée de l'adoption de la forme multidivisionnelle dans ces entreprises, nous renvoyons le lecteur à l'ouvrage de Chandler (1962).

[2] Chaque division correspond à un segment de marché et son équipe dirigeante dispose d'une autonomie totale concernant la gestion opérationnelle. Le rôle de la direction consiste à évaluer et coordonner les activités des divisions et prendre les orientations stratégiques de moyen et long terme.

évolution de la perception de l'organisation. En effet, il pose les bases de l'idée, qui sera développée plus tard par Simon (1951) et Williamson (1975), que les institutions facilitent les échanges et représentent une meilleure réponse que le marché à des contraintes contractuelles.

Cependant, les transactions sont également coûteuses au sein de l'entreprise. L'objectif de l'organisation interne de l'entreprise est donc de minimiser ces coûts. La nature de ces coûts de transaction peut être purement technologique ou alors peut correspondre à la nécessité d'inciter les employés à agir dans l'intérêt de l'entreprise. Ainsi, l'analyse théorique de l'organisation interne de l'entreprise s'articule autour de ces deux approches.

La première étudie la structure organisationnelle optimale, permettant de minimiser les coûts technologiques associés au traitement (ou l'acquisition) et à la transmission de l'information, et ceci dans un contexte où les intérêts des employés sont en concordance avec ceux de l'entreprise. Cette approche fournit une vision d'ensemble de l'organisation, et permet d'analyser les déterminants de sa structure, en termes de niveaux hiérarchiques, allocation de la prise de décision, etc.

La deuxième approche considère qu'une organisation est constituée d'individus dont les intérêts propres ne coïncident pas forcément avec ceux de leurs supérieurs ou collègues. Cette divergence pose un problème du fait de l'impossibilité de proposer aux employés des contrats suffisamment détaillés. Un premier obstacle à cela est l'asymétries d'information. En effet, au sein de l'entreprise les employés ont accès à des informations qu'ils sont les seuls à connaître (sélection adverse), et entreprennent des actions que l'employeur n'observe qu'imparfaitement (aléa moral). Le deuxième problème tient à l'impossibilité de proposer des contrats dont la mise en application peut être garantie par un tiers. Ainsi, l'objectif de l'employeur est de prendre les choix organisationnels et salariaux adéquats afin de motiver les employés à fournir les informations appropriées et à entreprendre les actions souhaitées.

Les quatre chapitres constituant cette thèse s'articulent autour des deux problématiques majeures concernant l'étude de l'organisation interne : la question de la structure optimale de l'organisation et la question de la rémunération incitative des employés. Dans les chapitres 1 et 2, nous étudions la structure organisationnelle optimale, en considérant que celle-ci résulte d'arbitrages d'ordre technologique mais aussi incitatif. Sur le plan théorique, cela nous permet de concilier les deux approches présentées plus haut. Nous supposons que pour motiver les employés, les employeurs utilisent des schémas de rémunération dont le coût dépend des conditions sur le marché du travail. Ainsi, nous analysons comment les organisations peuvent adapter leur structure à des évolutions institutionnelles ou en termes d'offre et de demande de main d'œuvre, ainsi que les répercussions de ces évolutions sur l'emploi et les salaires. La deuxième partie de la thèse est consacrée à une analyse plus approfondie des seuls schémas de rémunération. Dans le chapitre 3, nous étudions la structure de la rémunération optimale en présence de sélection adverse et aléa moral et lorsque l'exécution du contrat n'est pas garantie par un tiers. Cela nous permet de discuter les conditions sous lesquelles les contrats utilisés dans les chapitres 1 et 2 sont optimaux. De plus, nous apportons un éclairage nouveau sur l'utilisation simultanée par les entreprises de menaces de licenciement et de bonus informels pour motiver les employés à travailler. Enfin, dans le chapitre 4, nous étudions les effets de la possibilité pour les employés de se coordonner dans le choix de leurs efforts sur le contrat proposé par l'employeur.

1 La structure des organisations

Une vague de changements dans l'organisation du travail a marqué les dernières décennies. Les entreprises ont redéfini les limites de leurs activités, changé leur structure organisationnelle, réalloué la responsabilité et la prise de décision et modifié leurs systèmes d'information.

Osterman (1994) donne des indications concernant une première tendance. Il montre que les entreprises privilégient l'adoption de ce qu'il appelle des « formes de travail plus flexibles ». Près de la moitié des entreprises étudiées adoptent le travail en équipe, accompagné par un accroissement de la rotation de la main-d'œuvre et par l'introduction de plus en plus répandue de cercles de qualité (*quality circles*[3]).

La décentralisation de la prise de décision et la réduction des chaînes de commandement sont les autres évolutions relatées par les études empiriques. Elles ont pour conséquence l'apparition de formes organisationnelles plus plates, avec moins de niveaux hiérarchiques. L'étude de Rajan et Wulf (2006), basée sur 300 entreprises américaines sur la période 1986-1999, apporte des éléments confortant cette idée. Tout d'abord, ils montrent que le nombre de managers en contact direct avec le directeur général a augmenté : il est passé d'une moyenne de $4,4$ en 1986 à $7,2$ en 1999. De plus, le nombre de niveaux hiérarchiques qui séparent la direction générale de la base de l'organisation a connu une diminution de l'ordre de 25%.

Les répercussions de ces évolutions dépassent les limites de l'entreprise. En modifiant la façon de produire et de travailler, leurs effets s'étendent sur le fonctionnement de l'économie dans son ensemble. Il est donc important d'en identifier les causes et d'en prévoir les conséquences possibles.

1.1 La hiérarchie - organisation optimale

Les travaux qui étudient la structure optimale des organisations reposent essentiellement sur des considérations technologiques et négligent les problèmes d'incitation. Les agents se comportent comme des automates qui participent à la réalisation de l'objectif de l'entreprise. Dans un certain nombre de contributions récentes, l'arbitrage, qui permet d'expliquer la structure hiérarchique des organisations, est entre les coûts de traitement de l'information (ou acquisition des compétences) et les coûts de communication.

[3] Les *quality circles* sont des programmes dans lesquels les employés participent à la résolution de problèmes survenus ; voir Osterman (1994) p. 187.

Pour Radner (1992), Radner et Van Zandt (1992) ou encore Bolton et Dewatripont (1994), l'objectif de l'entreprise est de traiter (agréger) l'information qu'elle reçoit à intervalles réguliers le plus rapidement possible (chez Radner) ou en minimisant le temps total de travail des employés (chez Bolton et Dewatripont). Agréger les informations prend du temps (coût de traitement), traiter de l'information déjà agrégée par d'autres en prend également (coût de communication). S'il y a plusieurs personnes, elles peuvent se partager le travail, chacune traitant une fraction de l'information totale. Une fois le traitement fini, chaque employé transmet son agrégat à son supérieur et ainsi de suite jusqu'à ce que tous les agrégats soient réunis par une seule personne. Les auteurs définissent les conditions sous lesquelles la structure hiérarchique apparaît comme étant la forme organisationnelle optimale.

Outre la vision très mécanique du rôle des agents au sein de l'entreprise, ces théories réduisent l'activité de l'entreprise à une seule tâche, le traitement de l'information. Or, généralement le processus de production des entreprises englobe une multitude de tâches, nécessitant chacune des compétences spécifiques. Ainsi, le point de départ de Garicano (2000) est l'idée que l'organisation permet d'optimiser l'utilisation des compétences spécifiques des agents. Si l'acquisition de compétences est coûteuse, une fois acquises, elles restent attachées aux individus et peuvent donc être utilisées autant que possible dans les limites de leur temps de travail. Garicano suppose que l'entreprise dans son processus de production est confrontée à des tâches nécessitant différents niveaux de compétences, sans qu'il soit possible de déterminer *ex ante* la personne capable de résoudre chacune de ces tâches. Lorsqu'un client (ou un dossier) arrive, il est impossible de connaître le niveau de compétences que sa requête va mobiliser, avant de lui avoir consacré une partie de son temps. De plus, si l'employé n'est pas capable de donner satisfaction au client, il pourrait être dans l'incapacité de trouver qui exactement serait en mesure de s'en charger. En revanche, il peut toujours demander de l'aide à quelqu'un dans l'entreprise. Au sein de l'organisation, la communication est possible mais coûteuse en terme de temps. En effet, une personne peut apporter son aide à un nombre limité de collègues. De plus, pendant qu'elle les aide elle ne peut pas recevoir elle-même un client. Garicano montre que l'allocation efficace du travail consiste à utiliser le temps des plus compétents uniquement lorsque les personnes moins compétentes sont dans l'incapacité de résoudre un problème. Dans un contexte où il est coûteux de trouver la personne capable de résoudre un problème donné, l'organisation optimale est une hiérarchie basée sur les compétences (*knowledge based hierarchy*). Les agents au niveau le plus bas sont en charge des tâches les plus courantes. Ils reçoivent les problèmes et transmettent au niveau suivant ceux qu'ils ne sont pas en mesure de résoudre. Les agents aux niveaux supérieurs de la hiérarchie sont capables d'accomplir des tâches de plus en plus difficiles.

Le mode hiérarchique basé sur les compétences décrit le fonctionnement d'un large spectre d'organisations[4]. C'est le cas des services après-vente, des entreprises de services destinés aux professionnels ou encore de certains laboratoire de recherche, en particulier en sciences dures, où les assistants de recherche s'occupent des problèmes et tâches les plus simples et demandent de l'aide aux chercheurs confirmés lorsqu'ils sont dans l'impossibilité de trouver une réponse.

La structure de l'organisation dépend de l'arbitrage entre les coûts de communication et les coûts d'acquisition des compétences. Dans ce qui suit, nous allons voir dans quelle mesure des évolutions dans l'environnement économique affectent ces différents coûts et l'organisation de l'entreprise.

1.2 Organisation de l'entreprise et environnement économique

1.2.1 Changements technologiques et structure des organisations

Les dernières décennies ont été marquées par le développement des technologies de communication et information (TCI) et leur très large adoption au sein des

[4] Voir Garicano et Rossi-Hansberg (2006) pour des exemples supplémentaires.

entreprises. Aux États-Unis par exemple, la proportion d'employés utilisant un ordinateur[5] sur leur lieu de travail a doublé entre 1984 et 1993, atteignant presque 50%. En outre, alors qu'en 1977 seulement 38% des entreprises investissaient dans l'achat d'ordinateurs, elles étaient 87% à le faire en 1987. Ces changements ont affecté l'organisation du travail au sein des entreprises. Bresnahan, Brynjolfsson et Hitt (2002) trouvent, à partir de données d'entreprises, qu'il existe des complémentarités entre l'utilisation d'ordinateurs, l'emploi de davantage de main-d'œuvre qualifiée et la réorganisation du travail, marquée notamment par la décentralisation de la prise de décision au sein des entreprises. Par ailleurs, le changement technologique (biaisé en faveur de la main-d'œuvre qualifiée) est considéré comme un des éléments susceptibles d'expliquer l'accroissement des inégalités salariales au cours des années 1980-1990 aux États-Unis.

Pour comprendre ces évolutions Garicano et Rossi Hansberg (2006) proposent une analyse théorique des effets des TCI sur l'architecture des organisations et sur la structure des salaires. Ils construisent un modèle d'équilibre où les organisations sont des hiérarchies basées sur les compétences et où les agents sont hétérogènes dans leur coût d'acquisition des connaissances. Les TCI réduisent à la fois les coûts d'acquisition des connaissances et les coûts de communication. Ils montrent que la réduction du coût d'accès aux compétences permet d'expliquer les évolutions organisationnelles et salariales décrites plus haut. Ainsi, un coût d'acquisition des compétences plus faible se traduit par une décentralisation de la prise de décision au sein des entreprises. En effet, puisque les agents à tous les niveaux acquièrent plus de compétences, la probabilité qu'un problème soit résolu à un niveau plus faible de la hiérarchie augmente. L'inégalité salariale s'accroît alors à tous les niveaux.

L'effet d'une réduction du coût de communication est moins clair. Une telle réduction augmente le taux d'utilisation des connaissances des managers. En effet, ils peuvent traiter plus de tâches dans les limites de leur temps de travail. En

[5] Toutes les données citées dans ce paragraphe sont tirées de Lehr et Lichtenberg (1999).

conséquence, les managers acquièrent plus de compétences. L'effet sur les employés productifs est inverse, puisqu'il est moins coûteux de demander de l'aide à leurs supérieurs, ils peuvent acquérir moins de compétences. Ainsi, les inégalités salariales entre managers augmentent et celles entre employés diminuent.

1.2.2 Conditions sur le marché du travail et structure des organisations

Cependant, l'environnement économique et institutionnel auquel les entreprises font face est bien plus riche que les caractéristiques de la technologie d'information ou de communication. Des entreprises situées dans des pays avec un développement technologique similaire peuvent différer au niveau de leur taille ou structure. Kumar, Rajan et Zingales (1999) étudient les déterminants de la taille des entreprises dans des pays de l'OCDE. Ils montrent par exemple que les pays dotés d'un système juridique plus efficace abritent des entreprises dont la taille est plus importante. Davis et Henrekson (1999) comparent la distribution de l'emploi par industries et par taille des établissements en Suède et aux États-Unis. Il apparaît ainsi que le cadre institutionnel – système de taxation, accès au crédit ou encore régulations sur le marché du travail – affecte les choix industriels des entreprises ainsi que leurs tailles. Enfin, Caroli et Van Reenen (2001) montrent que la probabilité de procéder à un changement organisationnel (décentralisation et réduction des niveaux hiérarchiques) est négativement corrélée au coût relatif de la main-d'œuvre qualifiée.

Dans les Chapitres 1 et 2 de la thèse, nous proposons un modèle théorique permettant d'étudier comment la taille et la structure interne des entreprises sont affectées par des évolutions sur le marché du travail en termes d'offre et de demande mais aussi en terme d'intervention institutionnelle. Dans notre modèle, comme chez Garicano (2000), il est optimal d'organiser la production au sein de hiérarchies basées sur les compétences. Cependant, nous nous différencions de son analyse en considérant que les compétences des employés potentiels sont déjà acquises. Afin que les agents en fassent l'usage adéquat au sein de l'organisation, il est nécessaire de leur donner les incitations appropriées. Par conséquent, du point de vue de l'organisation, le coût des compétences est associé aux incitations à fournir aux employés.

Nous supposons qu'il existe une population de propriétaires d'entreprise qui sont les seuls à disposer d'une technologie permettant de produire[6]. Chacun d'entre eux a la possibilité de produire tout seul ou d'embaucher des employés pour l'aider. Les employés sont soit diplômés, soit non-diplômés et nous supposons qu'un agent diplômé est plus compétent que n'importe lequel des non-diplômés. Pour résoudre un problème, l'employé doit non seulement posséder des compétences, mais de plus il doit exécuter un effort. Nous considérons que le coût des incitations pour l'employeur dépend des conditions sur le marché du travail. En ce qui concerne la structure de l'organisation, nous conservons les hypothèses sous lesquelles Garicano montre que la hiérarchie basée sur les compétences est la forme organisationnelle optimale.

Dans le Chapitre 1, nous étudions des structures hiérarchiques à trois niveaux : une entreprise est composée de travailleurs non-diplômés, de cadres diplômés et d'un propriétaire de l'entreprise capable de résoudre les problèmes les plus difficiles. Dans ce chapitre, au sein de chaque population les agents sont hétérogènes et la productivité exacte de chacun est son information privée. Ainsi, le principal est confronté à deux problèmes informationnels : l'aléa moral et la sélection adverse. L'étude de la forme du contrat optimal n'est pas le propos du chapitre. Dans le Chapitre 3, nous nous concentrons sur l'analyse du contrat optimal dans un contexte similaire, lorsque la mise en oeuvre du contrat ne peut pas être garantie par un tiers. Dans le Chapitre 1, nous supposons que l'employeur propose un contrat par poste (travailleurs ou cadres) et que les incitations sont assurées par un salaire fixe accompagné d'un seuil d'exigence que l'employé doit atteindre afin de garder son poste. Ainsi, les conditions sur le marché du travail affectent à la fois le coût et la qualité de la main-d'œuvre. L'effet sur le coût est le suivant : si le niveau

[6] Les autres agents de l'économie ont la possibilité de travailler pour les propriétaires d'entreprise mais pas à leur propre compte.

de chômage diminue, un chômeur a plus de chances de retrouver du travail. Par conséquent, le licenciement n'est pas une menace forte et le salaire nécessaire pour motiver les employés est plus élevé. L'effet sur la productivité espérée passe par un processus de sélection *ex post*. En effet, à chaque période des employés (à l'équilibre peu compétents), qui n'ont pas réussi à atteindre l'objectif de performance fixé par l'employeur, sont renvoyés et remplacés. Ainsi, chaque distribution initiale, celle des non-diplômés et celle des diplômés, est partagée en deux distributions d'équilibre : la distribution des agents dans l'entreprise et celle des chômeurs. Sur un marché où le taux de chômage est faible, les employés sont moins sélectionnés (leur distribution est plus proche de la distribution initiale) et leur productivité espérée est donc plus faible. Les décisions de chaque principal concernant le nombre d'agents à engager aux différents niveaux hiérarchiques et l'allocation des tâches entre employés et managers dépendent des coûts et des productivités relatifs de la main-d'œuvre, et donc varient avec les conditions (d'offre et de demande mais aussi institutionnelles) sur le marché du travail. Ces choix affectent en retour la structure d'équilibre de l'emploi et des salaires.

La structure du modèle du Chapitre 2 est très proche de celle que nous venons de développer. La seule différence consiste dans le fait que, dans ce chapitre, au sein de chaque population les agents sont homogènes. Nous étudions le choix du nombre de niveaux hiérarchiques en fonction des caractéristiques de l'offre de main-d'œuvre, ainsi que les conséquences d'un éventuel changement organisationnel sur l'emploi et les salaires. Plus précisément, nous comparons les deux structures suivantes : une hiérarchie à trois niveaux, où l'employeur embauche à la fois des diplômés et des non-diplômés et une hiérarchie à deux niveaux, où seuls des diplômés sont engagés. Ainsi, le niveau en dessous du propriétaire est toujours occupé par des diplômés. Ajouter un niveau de non-diplômés permet au principal de réduire l'ensemble des tâches dévolues aux qualifiés, dans la mesure où ces derniers ne sont alors sollicités que pour les problèmes que leurs subordonnés n'ont pas su résoudre. Ceci a deux conséquences :

tout d'abord le nombre de qualifiés diminue ; de plus, puisque leur charge de travail est réduite, leur salaire incitatif est plus faible. La décision d'adopter une hiérarchie à trois niveaux résulte donc de l'arbitrage entre la réduction du coût salarial des diplômés et le nouveau coût salarial dû à l'embauche des non-diplômés. Ainsi, dans la mesure où l'augmentation de l'offre de main-d'œuvre qualifiée réduit leur option de sortie et donc leur salaire, elle rend moins avantageuse l'adoption d'une structure hiérarchique à trois niveaux. Cela se traduit par une modification qualitative des emplois destinés aux diplômés. En effet, ceux-ci sont davantage sollicités, même pour des tâches simples, ce qui induit une hausse de leur salaire. Ainsi, nous montrons qu'une offre plus importante d'agents diplômés peut conduire à un aplatissement des organisations, qui s'accompagne par l'augmentation de la demande pour des employés qualifiés et par l'accroissement de l'écart salarial entre diplômés et non-diplômés.

2 La rémunération incitative

La plupart des travaux théoriques étudiant l'organisation interne de l'entreprise s'intéressent aux conflits d'intérêts existants en son sein, et à la manière dont la rémunération et l'organisation devraient être structurés afin de les résoudre. Les Chapitres 3 et 4 étudient les schémas de rémunération incitatifs[7]. La rémunération proposée à l'employé est fixée dans le contrat de travail dont les caractéristiques dépendent de la possibilité de s'engager *ex ante* (au moment de la signature) sur une information vérifiable *ex post* (au moment de la mise en application) par une tierce partie (cour de justice par exemple). L'enjeu porte sur la capacité de chacune des parties de prouver le cas échéant que le contrat n'a pas été respecté afin que la juridiction compétente puisse imposer les sanctions correspondantes. Dans le cas contraire, le respect du contrat n'est pas garanti. En ce sens, il existe des cas où la

[7] Pour une revue de la littérature plus exhaustive, voir Gibbons (1998), Gibbons et Waldman (1999) et Prendergast (1999).

performance des employés est observable au sein de l'entreprise mais invérifiable par un tiers. La possibilité de vérifier ou non l'information observée définit le caractère explicite ou implicite des clauses du contrat. Les clauses explicites sont basées sur des éléments vérifiables, inscrits dans le contrat. Les clauses implicites relèvent d'un contrat relationnel (*relational contract*) entre les parties concernées. Ce dernier repose sur des promesses faites par les cocontractants et son respect est garanti par la possibilité pour chacun d'entre eux d'interrompre la coopération, si l'autre n'honore pas ses promesses.

Dans l'analyse des relations employeurs-employés en présence d'aléa moral, une grande place a été accordée aux contrats explicites. Dans ce contexte, le modèle principal-agent standard pose le cadre de référence. Considérons, ainsi, un employeur (principal) qui embauche un employé (agent) pour la réalisation d'une certaine tâche (production). La seule information observable pour le principal et vérifiable par un tiers est la production réalisée. Cette dernière croît avec l'effort de l'agent mais peut comporter également une part aléatoire. En effectuant peu d'effort l'employé pourrait réaliser une haute performance parce qu'il a eu de la chance, et inversement obtenir une mauvaise production malgré un effort élevé. L'objectif du principal est de donner à l'employé les incitations adéquates afin que le niveau d'effort désiré soit effectué. Le caractère plus ou moins incitatif de la rémunération dépend de la façon dont elle varie avec la performance réalisée, et *a fortiori* avec l'effort fourni par l'employé. L'agent aura les « bonnes » incitations à travailler s'il supporte totalement les conséquences de ses choix (coûts et bénéfices). On dit alors que l'agent est demandeur résiduel. La rémunération doit donc être structurée de telle façon que l'employé tienne compte de l'impact de l'effort choisi sur le niveau des bénéfices réalisés par l'entreprise. Cependant, si l'employé est averse au risque, ceci n'est pas toujours souhaitable. En effet, une rémunération basée sur la performance augmente le risque supporté par l'agent. Or, lui faire supporter du risque est coûteux pour le principal puisque *in fine* ce dernier doit le compenser pour le risque encouru. Dans ce cas, le coût de l'effort pour l'employeur est plus élevé, donc le niveau d'effort qui est fourni par l'agent est plus faible que celui du premier rang.

Les Chapitres 3 et 4 de cette thèse étendent l'analyse standard en deux directions. Tout d'abord, dans la mesure où une grande partie des moyens de motivation utilisés dans les entreprises repose sur des promesses informelles et des conventions tacites, nous nous intéresserons aux moyens d'incitation implicites. Ensuite, nous aborderons le travail des employés comme un ensemble d'activités inter-dépendantes et nous étudierons l'impact de la possibilité pour les agents à se superviser mutuellement sur le contrat proposé par le principal.

2.1 Les contrats de travail implicites

Les contrats de travail utilisés par les entreprises semblent moins formalisés que ce que la théorie de l'agence le suggère. En effet, dans de nombreux cas, les bonus ou les promotions reposent sur des promesses informelles plutôt que sur des critères de performance clairement inscrits dans le contrat de travail. Chez Lincoln Electric par exemple[8], une partie de la rémunération est fonction de la production réalisée par l'employé, mais environ la moitié de cette rémunération correspond à des bonus basés sur l'évaluation subjective de sa capacité à coopérer, à innover, de son autonomie, etc. MacLeod et Parent (1997) soulignent que 14% des employés américains interrogés[9] (toutes industries confondues) déclarent recevoir des bonus ou promotions fondées sur des promesses informelles. Hayes et Schaefer (2000) fournissent également des éléments empiriques qui témoignent de l'existence de bonus basés sur des critères observables uniquement par les parties participant au contrat. Ils étudient la rémunération de cadres dirigeants accordée par les conseils d'administration. Ils montrent que des variations dans cette rémunération, non

[8] Exemple tiré de Gibbons (1998).
[9] L'étude est basée sur la National Longitudinal Survey of Youth (NLSY) de 1988-1990.

expliquées par les résultats actuels des entreprises, prédisent des variations futures de leurs performances.

L'analyse des contrats relationnels nous permet d'avoir une meilleure compréhension des incitations utilisées au sein des entreprises mais également entre elles[10]. Ce type de contrats est mis en œuvre non pas par une cour de justice mais par la crainte qu'à l'avenir une des parties mette fin à la coopération. En ce sens ils sont auto-exécutoires (*self-enforcing*). En effet, chacun respecte le contrat afin d'éviter d'être « puni » par l'action de(s) autre(s) personne(s) concernée(s). Dans le cadre des relations employeur-employé, la punition exercée par l'employé consiste à arrêter de travailler ou à quitter l'entreprise qui n'honore pas ses promesses salariales alors que l'employeur peut licencier ou réduire la rémunération de l'agent qui n'a pas réalisé la performance convenue.

En plus de la possibilité de punir la partie qui ne respecte pas sa part du contrat relationnel, la perception de chacun sur ces punitions éventuelles est cruciale pour garantir le respect de l'arrangement. Blinder et Choi (1990) étudient cette perception à l'aide d'une enquête menée auprès d'un petit échantillon de managers d'entreprises américaines. D'après leur étude, 95% des managers interrogés estiment qu'une réduction des salaires jugée « injuste » par les employés conduit ces derniers à réduire leur effort, et 85% estiment qu'une telle réduction entraînerait une augmentation des départs.

D'un autre côté, si l'employé n'atteint pas la performance convenue, alors 90% des managers déclarent que l'employé en question doit être renvoyé. Dans une étude également basée sur des enquêtes, Bewley (1999) rapporte que la majeure partie des managers déclare utiliser la performance comme critère pour licencier ou non une personne. Même si la personne n'est pas renvoyée immédiatement, une mauvaise performance augmente la probabilité qu'elle le soit à la prochaine vague

[10] L'utilisation des contrats implicites ne se limite pas aux relations au sein de l'entreprise. Pour une analyse plus complète et notamment pour d'autres applications de ce type de contrats voir MacLeod (2007).

de licenciements.

Nous proposons à présent une brève revue de ce que nous enseignent les modèles théoriques existants sur la structure de la rémunération dans le cas où la performance des employés n'est pas vérifiable, mais est observable[11] par toutes les parties participant au contrat. Nous discutons également les apports et les résultats obtenus dans le Chapitre 3 de cette thèse.

2.1.1 Information symétrique.

Une large partie des travaux sur les contrats implicites au sein de l'entreprise, supposent que l'information est symétrique. Pour présenter cette branche de la littérature, nous nous basons sur le travail de MacLeod et Malcomson (1989). Ils s'intéressent à la relation entre un employeur et un employé, tous deux neutres au risque et vivant un nombre infini de périodes. Comme précédemment, la production nécessite un effort de la part de l'employé. Cependant, cet effort est supposé observable dans l'entreprise, mais impossible à vérifier par un tiers[12]. Si les deux parties n'arrivent pas à se mettre d'accord sur les termes du contrat ou si la relation contractuelle est interrompue, chacune reçoit l'utilité associée à son option de sortie : son indemnité de chômage pour l'employé, par exemple et le niveau de production sans employé pour l'employeur. Si l'agent accepte le contrat et effectue l'effort alors la production est réalisée et génère un certain surplus. Les deux parties étant neutres par rapport au risque, le salaire correspond à un transfert. Dans ce cadre, MacLeod et Malcomson montrent que si le surplus créé par la relation entre le principal et l'agent est suffisamment important par rapport à la somme des options de sortie, alors il est toujours possible de trouver un transfert qui permet de garantir les

[11] Voir MacLeod (2003) et Levin (2003) pour des contrats relationnels basés sur une évaluation subjective observable uniquement par une des parties prenant part au contrat.

[12] De plus, aucune évaluation vérifiable de la performance de l'employé n'est disponible.

incitations des deux parties[13]. En effet, dans la mesure où la relation contractuelle est mutuellement profitable et où la déviation d'une des deux parties conduit à l'interruption du contrat, il est dans l'intérêt commun de l'employeur et de l'employé de se mettre d'accord sur un transfert garantissant la pérennité de leur relation.

La forme du contrat d'équilibre peut dépendre des conséquences de l'action de la partie lésée sur la réputation de la partie ayant déviée et donc sur l'option de sortie de cette dernière. Pour illustrer notre propos, considérons les modèles de Shapiro et Stiglitz (1984) et Bull (1987). Shapiro et Stiglitz considèrent un cadre où le marché du travail est anonyme. Si le principal ne paie pas un bonus basé sur la performance, l'employé peut le punir en interrompant son contrat de travail. L'option de sortie de l'employeur correspond au profit qu'il pourra obtenir en embauchant une nouvelle personne sur le marché du travail. Le bonus n'est donc pas crédible et, pour inciter les employés, le principal propose un salaire fixe accompagné de la menace de renvoyer l'employé s'il ne fait pas l'effort. Dans le modèle de Bull (1987), en revanche, il n'y a pas d'anonymat. Ainsi, un employeur qui dévie perd sa réputation sur le marché de la main-d'œuvre. Son option de sortie correspond à son profit s'il est le seul à travailler. Dès lors, garder un employé est une source de rente pour le principal et le paiement d'un bonus peut être crédible.

MacLeod et Malcomson (1998) montrent qu'à structure informationnelle donnée (marché du travail anonyme), la nature du contrat relationnel d'équilibre dépend des conditions sur le marché du travail. Lorsqu'il y a du chômage et que les employés sont homogènes, on retrouve le cadre de Shapiro et Stiglitz, et le seul contrat incitatif auto-exécutoire est le salaire fixe accompagné de la menace de licenciement. En effet, il est impossible pour l'employeur de promettre de façon crédible de payer un transfert conditionnel à l'effort fourni, puisqu'une fois le travail accompli il peut

[13] Levin (2003) obtient la même condition dans un contexte avec information asymétrique. Cependant, l'asymétrie de l'information porte sur la période actuelle et l'information des deux parties est symétriques quant à la valeur de continuation espérée. Or, c'est le surplus espéré associé à la continuation de la relation contractuelle qui compte pour la crédibilité des transferts.

licencier l'employé et le remplacer sans coût par un chômeur. En revanche, si l'offre de main-d'œuvre est insuffisante, le paiement d'un bonus basé sur la performance peut être un contrat relationnel crédible. En effet, dans cette nouvelle configuration, garder un employé est une source de rente pour l'employeur. Il préfère rémunérer l'agent plutôt que de le voir partir.

2.1.2 Information asymétrique et sélection *ex post*

Les modèles présentés ci-dessus permettent donc d'identifier les conditions sous lesquelles le contrat de travail prend la forme d'un salaire fixe et une menace de licenciement ou alors d'un bonus payé en fin de période et basé sur la performance. Dans le Chapitre 3, nous caractérisons le contrat relationnel optimal lorsque l'information est asymétrique et l'asymétrie porte à la fois sur une caractéristique des agents et sur leur effort. Ainsi, nous obtenons des conditions sous lesquelles les deux formes d'incitation, bonus et licenciement, sont conjointement utilisées. Nous supposons que les employés potentiels sont hétérogènes et la productivité exacte de chacun est son information privée. Levin (2003) étudie, également, le contrat relationnel optimal en présence d'information cachée du côté de l'agent. Toutefois, dans son modèle, l'asymétrie d'information porte sur une caractéristique attribuée à chaque début de période de façon aléatoire et indépendante de ce qui a pu se produire dans le passé. Il analyse les restrictions imposées par le caractère implicite du contrat sur la possibilité de révéler l'information de l'agent. Lorsque, comme dans notre modèle, l'asymétrie d'information porte sur une caractéristique persistante dans le temps, le caractère dynamique de la sélection adverse et l'impossibilité de s'engager sur un contrat de long terme augmentent le coût de la révélation *ex ante*, pouvant la rendre sous-optimale. Nous restreignons donc notre analyse, dans le Chapitre 3, au cas où l'employeur propose un contrat unique à tous les employés embauchés.

Nous supposons qu'à chaque période l'employé est confronté à des tâches de difficultés variables, qu'il peut réussir s'il fait un effort coûteux et si sa productivité est suffisamment élevée. Plus une tâche est difficile et plus la productivité nécessaire pour son succès est élevée. L'objectif du principal est d'inciter les employés à fournir de l'effort sur le plus large ensemble possible de tâches. A la fin de chaque période, le principal observe si la tâche a été réussie ou non ainsi que sa difficulté. Dès lors, réussir une tâche difficile permet de signaler sa haute productivité. Afin d'exposer nos principaux résultats, supposons pour le moment que la productivité espérée des employés est plus élevée que celle des chômeurs. Ceci est un résultat endogène du modèle que l'on discute par la suite.

Nous montrons tout d'abord que, même en présence de chômage, le bonus peut être un moyen d'incitation crédible pour certaines tâches. Si une tâche réussie permet à l'employé de signaler que sa productivité est plus élevée que celle du chômeur moyen, il existe un gain pour le principal à garder l'agent. Dans le cas où ce gain est « suffisamment » élevé, le paiement d'un bonus devient crédible. Nous montrons que ceci est possible lorsque la distribution initiale des agents est « suffisamment » hétérogène. Dès lors, le bonus peut être utilisé pour motiver les employés à effectuer des tâches difficiles. En plus du bonus, le principal peut utiliser un salaire fixe et une menace de licenciement afin d'inciter les employés à réussir les tâches simples. En effet, il propose un contrat relationnel qui spécifie le salaire fixe et un seuil de performance minimale que l'agent doit atteindre afin de garder son poste.

Nous montrons également que, s'il est impossible d'utiliser un bonus, la menace de licenciement en cas de performance insuffisante n'est pas crédible pour motiver les employés à résoudre des tâches difficiles. En effet, si le seuil d'exigence est très élevé, alors la productivité espérée d'un employé qui n'arrive pas à l'atteindre est supérieure à celle d'un chômeur. Dans ce cas l'employeur préfère garder l'employé plutôt que de le licencier. En anticipant cela, l'agent ne fait pas l'effort requis pour atteindre le seuil d'exigence en question.

L'écart de productivité entre employés et chômeurs que nous avons supposé plus haut est crucial pour nos résultats. Dans le modèle cet écart est obtenu de façon endogène. Le mécanisme de sélection *ex post* est identique à celui décrit dans la sous-section 1.2.2. À chaque période des employés peu productifs sont renvoyés et remplacés. Ainsi, à l'équilibre stationnaire, nous obtenons deux distributions de productivité endogènes : sur le marché du travail et dans l'entreprise. Elles vérifient la propriété supposée plus haut, à savoir la productivité espérée des employés excède celle des chômeurs. L'écart entre ces deux distributions dépend des conditions sur le marché du travail, taux de chômage et taux de rotation exogène et affecte à son tour la composition du contrat optimal.

2.2 Le travail en équipe

Comme le montre Osterman (1994) l'adoption accrue du travail en équipe est une des transformations de l'organisation du travail dans les entreprises. Nous développons à présent les arbitrages relatifs à son adoption et présentons les apports du dernier chapitre de la thèse qui propose une justification théorique de la complémentarité entre le travail en équipe et certaines pratiques managériales, documentée par les études empiriques.

2.2.1 L'arbitrage entre coopération et incitations

Faire travailler les employés ensemble à la réalisation d'un projet peut être source de gains de productivité. De même, comme le souligne Itoh (1991), l'aide entre collègues peut générer des synergies et améliorer la performance de chacun. Pour bénéficier de ces gains, il faut inciter les employés à s'impliquer dans le projet commun ou à s'entraider. Ainsi, la rémunération doit être basée sur la performance commune dans le premier cas, ou doit dépendre positivement de la performance du collègue dans le second.

Cependant, ces gains de productivité ont deux contreparties. Dans le modèle de Itoh (1991), il s'agit d'un accroissement du coût de l'effort du fait de la perte de spécialisation de l'agent. Le deuxième effet, qui lui est purement incitatif, est lié au fait qu'il est plus coûteux d'inciter un agent lorsque la mesure utilisée pour baser sa rémunération reflète avec moins de précision l'effort fourni. Pour illustrer le problème incitatif associé au travail en équipe, considérons le cadre de Holmström (1982) et supposons que plusieurs personnes participent à la réalisation d'une tâche et que la production obtenue reflète fidèlement (sans aléa) la somme des efforts fournis par chacun. Le profit associé à la production réalisée est partagé entre les membres de l'équipe. Dans ce cas, l'effort additionnel de chacun est à l'origine d'une augmentation de la production (espérée) partagée par tous. Chaque agent décide du niveau d'effort qu'il veut fournir sans tenir compte du bénéfice (de l'externalité positive) que son effort additionnel génère pour les autres membres de l'équipe. Par conséquent, le niveau d'effort individuellement choisi est inférieur au niveau qui maximiserait l'utilité jointe de tous les membres de l'équipe.

2.2.2 Supervision mutuelle et coût des incitations

Afin d'obtenir la coopération des agents et de les faire travailler en équipe, une rémunération basée sur la performance du groupe est nécessaire. Cependant, accompagner le travail en équipe par d'autres pratiques managériales peut également être à l'origine de gains de productivité. Appelbaum et Batt (1994) évoquent, par exemple, l'autonomie accordée aux employés dans la gestion des ressources humaines. Dans les entreprises étudiées par les auteurs, les membres des équipes ont la possibilité de fixer des règles disciplinaires et de participent à la sélection des nouvelles recrues. L'existence de complémentarités entre l'adoption du travail en équipe et la mise en place de certaines pratiques managériales est également documentée par Ichniowski, Pennushi et Shaw (1997). Ces auteurs montrent que le travail en équipe conduit à l'accroissement de la productivité lorsque les employés ont la possibilité de gérer les problèmes au quotidien et lorsque des pratiques améliorant la supervision mutuelle sont mises en place.

Dans le Chapitre 4, nous étudions dans quelle mesure la possibilité pour les employés d'interagir, de décider de normes d'effort, de se superviser mutuellement et de se punir en cas de non respect de ces normes, affecte le contrat incitatif proposé par le principal. Nous considérons le cadre suivant. Un employeur embauche deux employés, neutres par rapport au risque mais protégés par une clause de responsabilité limitée, pour la réalisation d'un projet commun. Ce dernier peut se solder soit par une réussite, soit par un échec et le principal observe uniquement le résultat. Le succès du projet dépend positivement des efforts des employés. Nous supposons qu'au sein de l'équipe les agents observent des signaux sur l'effort choisi par chacun. De plus, les coéquipiers peuvent s'entendre[14] sur les termes d'un contrat basé sur toute l'information qu'ils observent. Dans notre modèle, les employés décident du niveau d'effort à fournir et s'engagent sur des transferts à payer au coéquipier en cas de mauvais signal. Chaque employé a donc une double motivation pour travailler : toucher le bonus associé à une bonne performance et éviter d'être « puni » par son collègue. Par conséquent, pour obtenir un niveau d'effort donné, l'employeur peut proposer un bonus plus faible. Cependant, étant protégés par la responsabilité limitée, les employés ne peuvent pas s'engager sur n'importe quel niveau de transfert. En effet, dans le pire des cas l'agent touche le salaire faible (parce que le projet a échoué) et doit payer un transfert à son collègue (car son signal a été mauvais). Si la somme des transferts est trop élevée (au delà d'un seuil exogène), alors nous supposons que l'employé peut quitter l'entreprise sans en effectuer le paiement. Pour que les coéquipiers puissent s'engager sur une punition, il faut que le principal augmente le salaire payé en cas de mauvaise performance. Le fait que les agents puissent se superviser mutuellement permet de réduire le bonus payé en cas de succès, mais demande en contrepartie que le salaire payé en cas d'échec soit plus élevé. Il est profitable pour le principal de laisser les employés

[14] L'idée sous-jacente est que des règles tacites et des échanges en cas de non respect de ces règles peuvent se mettre en place entre agents qui ont des contacts réguliers. Cependant, nous supposons que le contrat peut être mis en oeuvre, sans étudier en détails le mécanisme garantissant sa mise en application.

signer un contrat entre eux, lorsque ce dernier permet de réduire fortement le bonus. Nous montrons que cela arrive lorsque les signaux, observés au sein de l'équipe, sur l'effort de chacun sont suffisamment informatifs. La forme du contrat, proposé par le principal, dépend également de la contrainte de responsabilité limitée. Si celle-ci n'est pas très contraignante, alors l'employeur préfère utiliser dans une moindre mesure la possibilité de coopération des agents et proposer dans ce cas un bonus plus élevé en cas de succès et un salaire plus faible en cas d'échec du projet.

Puisqu'en facilitant l'observabilité de l'effort des employés au sein de l'équipe le principal peut obtenir le même niveau d'effort à un coût salarial plus faible, il peut être dans son intérêt de mettre à la disposition des agents une technologie de supervision, même si cette dernière nécessite un effort coûteux de leur part. Dans ce contexte, nous comparons deux formes de supervision au sein de l'équipe : la supervision mutuelle, où chacun des agents supervise son collègue, et la supervision unilatérale, où le principal délègue à un des employés la technologie de supervision et la charge de proposer un contrat à l'autre employé. Nous montrons que la supervision mutuelle est préférable à la supervision unilatérale lorsque l'effort de supervision est plus coûteux.

Dans cette thèse nous étudions le lien entre les choix organisationnels internes à l'entreprise et son environnement économique. Nous montrons qu'aussi bien les possibilités d'interactions entre les employés que les conditions sur le marché du travail affectent les schémas d'incitation utilisés afin de motiver les employés. De plus, en mettant l'accent sur le lien entre le coût des incitations et les conditions sur le marché du travail, nous analysons les répercussions d'évolutions dans ces conditions sur les choix organisationnels des entreprises.

Bibliographie

Appelbaum E. and R. Batt (1994), *The New American Workplace*, Cornell University Press

Bewley T. (1999), *Why Wages Don't Fall During a Recession*, Harvard University Press

Blinder A. and D. Choi (1990), « A Shred of Evidence on Theories of Wage Stickiness », *Quarterly Journal of Economics*, vol. 55, pp. 1003-1015.

Bolton P. and M. Dewatripont (1994), « The Firm as a Communication Network », *Quarterly Journal of Economics*, vol. 109, pp. 809-839.

Bresnahan T., E. Brynjolfsson and L. Hitt (2002), « Information Technology, Workplace Organization and the Demand for Skilled Labor : Firm-level Evidence », *Quarterly Journal of Economics*, vol. 117, pp.339-376

Bull C. (1987), « The Existence of Self-Enforcing Implicit Contracts », *Quarterly Journal of Economics*, vol. 102, pp. 147-159.

Caroli E. and J. Van Reenen (2001), « Skill Biased Organizational Change ? Evidence from a panel of British and French Establishments », *Quarterly Journal of Economics*, vol. 116, pp. 1449-1492.

Chandler A. (1962), « Strategy and Structure », Cambridge, MA : MIT Press

Coase R. (1937), « The Nature of the Firm », *Economica*, vol. 4, pp. 386-405.

Davis S. and M. Henrekson (1999), "Explaining National Differences in the Size and Industry Distribution of Employment", *Small Business Economics*, vol. 12, pp. 59-83.

Garicano L. (2000), « Hierarchies and the Organization of Knowledge in Production », *The Journal of Political Economy*, vol. 108, pp. 874-904.

Garicano L. and E. Rossi-Hansberg (2006), « Organization and Inequality in a Knowledge Economy », *Quarterly Journal of Economics*, vol. 121, pp. 1383-1436.

Gibbons R. (1998), « Incentives in Organizations », *The Journal of Economic Perspectives*, Vol. 12, pp. 115-132

Gibbons R. and M. Waldman (1999), « Careers in organizations : Theory and evidence », *Handbook of Labor Economics*, in : O. Ashenfelter & D. Card (ed.), vol. 3, pp. 2373-2437

Hayes R. and S. Schaefer, 2000, "Implicit Contracts and the Explanatory Power of Top Executive Compensation for Future Performance," *RAND Journal of Economics*, Vol. 31, pp. 273-293.

Holmström B. (1982), « Moral Hazard in Teams », *The Bell Journal of Economics*, vol. 13, pp. 324-340

Ichniowski C., G. Prennushi and K. Shaw (1997), « The Effects of Human ressources Management Practices on Productivity », *American economic Review*, vol. 59, pp. 291-313.

Itoh H. (1991), « Incentives to Help in Multi-agents Situations », *Econometrica*, vol. 59, pp. 611-636

Kumar K., R. Rajan and L. Zingales (1999), « What Determines firm size ? », NBER working paper 7208.

Lehr B. and F. Lichtenberg (1999), « Information Technology and its Impact on Productivity : Firm-level Evidence from Government and Private Data Sources, 1977-1993 », *Canadian Journal of Economics*, vol. 32, pp. 335-362.

Levin J. (2003), « Relational Incentive Contracts », *American Economic Review*, vol. 93, pp. 835-847.

MacLeod W.B. (2003), « Optimal Contracting with Subjective Evaluation », *American Economic Review*, vol. 93, pp. 216-240.

MacLeod W.B. (2007), « Reputations, Relationships and the Enforcement of Incomplete Contracts » Forthcoming in *Journal of Economic Literature*.

MacLeod W.B. and J.M. Malcomson (1989), « Implicit Contracts, Incentive Compatibility, and Involuntary Unemployment », *Econometrica*, vol. 57, pp. 447-480.

MacLeod W.B. and J.M. Malcomson (1998), « Motivation and Markets », *American Economic Review*, vol. 88, pp. 388-411.

MacLeod W.B. and D. Parent (1997), « Jobs Characteristics and the Form of Compensation », Mimeo, University of Southern California.

Osterman P. (1994), « How Common is Workplace Transformation and Who adopts It ? », *Industrial and Labor Relations Review*, Vol. 47(2), pp. 173-188.

Prendergast C. (1999), « The Provision of Incentives in Firms » *Journal of Economic Literature*, Vol. 37, pp. 7-63.

Radner R. (1992), « Information Processing in Firms and Returns to scale », *Journal of Economic Literature*, vol. 30, pp. 1382-1415.

Radner R. and T. Van Zandt (1992), « Information Processing in Firms and Returns to scale », *Annales d'Économie et de Statistique*, vol. 25-26, pp. 265-298.

Rajan R. and J. Wulf (2006), « The Flattening Firm : Evidence from Panel Data on the Changing Nature of Corporate Hierarchies », *The Review of Economics and Statistics*, vol. 88, pp. 759-773.

Shapiro, C. and J. Stiglitz (1984), « Equilibrium Unemployment as a Worker Discipline Device », *American Economic Review*, vol. 74, pp. 433-444.

Simon H. (1951), « A Formal Theory of the Employment Relationship », *Econometrica*, vol. 19, pp. 293-305.

Simon H. (1991), « Organizations and Markets », *Journal of Economic Perspectives*, vol. 5, pp. 25-44.

Williamson O. (1975), « Markets and Hierarchies : Analysis and Antitrust Implications », New York : Free Press

Chapitre 1

Incitations, Marchés et Hiérarchies basées sur les Compétences

1 Introduction

Une vague de changements dans l'organisation du travail a marqué les dernières décennies. Les entreprises ont redéfini les limites de leurs activités, changé leur structure organisationnelle et réalloué la prise de décision[1]. En modifiant la façon de produire et de travailler, les répercussions de ces évolutions dépassent les limites de l'entreprise et peuvent toucher au fonctionnement du marché du travail[2]. Il nous semble donc important d'en identifier les causes et d'en prévoir les conséquences possibles.

La majeure partie des modèles théoriques consacrés à l'analyse de l'organisation interne des entreprises a porté sur les conflits d'intérêts et la façon de les résoudre

[1] Rajan et Wulf (2006) montrent, à partir de données américaines sur la période 1986-1995, que les entreprises étudiées comportent moins de niveaux hiérarchiques. Voir également Appelbaum et Batt (1994) et Roberts (2004) pour des exemples.

[2] Caroli et van Reenen (2001), par exemple, montrent que les changements organisationnels, adoptés par les entreprises françaises et anglaises étudiées, sont biaisés en faveur de la main-d'œuvre qualifiée.

par le biais de contrats[3] (explicites ou implicites) ou en choisissant la forme organisationnelle appropriée[4]. Or, ces modèles apportent peu d'éclairage sur les forces à l'origine de la structure des organisations et de leur transformation. Une branche récente de la littérature sur les hiérarchies endogènes apporte néanmoins des éléments de réponse à ces questions. Les travaux qui étudient la structure optimale des organisations reposent essentiellement sur des considérations technologiques et négligent les problèmes d'incitation, en supposant que les intérêts des agents sont parfaitement alignés avec ceux de l'organisation. Pour Radner (1992,1993), Radner et Van Zandt (1992) et Bolton et Dewatripont (1994), l'objectif de l'entreprise est de traiter (agréger) l'information qu'elle reçoit à intervalles réguliers le plus rapidement possible (chez Radner) ou en minimisant le temps total de travail des employés (chez Bolton et Dewatripont). Ainsi, dans ces modèles, l'organisation optimale résulte de l'arbitrage entre les coûts de traitement et les coûts de transmission de l'information. Ces auteurs définissent les conditions sous lesquelles la structure hiérarchique est la forme organisationnelle optimale. Dans ces articles, l'activité de l'entreprise se présente sous la forme d'une seule tâche, le traitement de l'information. Garicano (2000) enrichit cette approche en considérant que l'entreprise est confrontée au traitement de problèmes de difficultés variées et l'objectif de l'organisation est de coordonner l'activité de travailleurs spécialisés. Il considère que la compétence est un facteur essentiel à la production et les agents peuvent acquérir différents niveaux de compétence. Au sein de l'entreprise un employé incapable de résoudre un problème peut demander de l'aide à un de ses collègues. Cependant, une telle communication est coûteuse en termes de temps et le temps de chacun est limité. Ainsi, le principal arbitrage pour l'organisation se fait entre les coûts d'acquisition des compétences et le coût de la communication. Garicano montre que dans ce contexte, la forme organisationnelle optimale est une hiérarchie basée sur les compétences (*knowledge*

[3] Prendergast (1997) et Gibbons (1998) proposent des revues de cette littérature.
[4] Calvo et Wellisz (1978), Aghion et Tirole (1997) parmi d'autres.

based hierarchy). Dans une telle organisation, les agents du niveau le plus bas sont en charge des problèmes les plus simples, alors que les cadres des niveaux supérieurs se consacrent uniquement aux problèmes trop difficiles pour leurs subalternes. Ainsi, la compétence des personnes les plus aptes est efficacement utilisée puisqu'elle est sollicitée uniquement pour des problèmes que les personnes moins compétentes n'ont pas su résoudre. L'article de Garicano et Rossi-Hansberg (2006) étend cette analyse afin d'étudier les effets de changements technologiques sur la structure des organisations et les salaires. Les auteurs considèrent une économie avec des agents hétérogènes dans leur capacité à acquérir des compétences. La production est organisée au sein de hiérarchies basées sur les compétences. Garicano et Rossi-Hansberg étudient l'appariement d'agents de capacités différentes au sein des entreprises et la structure d'équilibre des revenus. Dans la mesure où les intérêts des agents sont toujours alignés avec ceux de l'organisation et les seuls coûts d'acquisition ou communication de l'information sont ceux de nature technologique, leur analyse se concentre sur les répercussions de changements dans ces coûts sur la structure des organisations et des salaires. Le rôle crucial des évolutions technologiques[5] récentes sur la transformation du travail au sein des entreprises est incontestable[6]. Néanmoins, elles ne peuvent pas tout expliquer, puisque les entreprises diffèrent par leur taille (Kumar, Rajan et Zingales (1999)) ou encore en termes de nombre d'employés par cadre (Acemoglu et Newman (2002)), entre pays dont l'accès aux technologies de l'information est proche.

D'autres aspects de l'environnement économique de l'entreprise peuvent ainsi influer sur sa structure organisationnelle. Notre analyse porte sur l'impact d'évolutions sur le marché du travail. Plus particulièrement, nous nous intéressons à l'effet de changements réglementaires ou de l'offre de la main-d'œuvre sur la structure des

[5] Le développement des outils informatiques ou encore d'internet.

[6] Bresnahan et al. (2002), parmi d'autres, montrent que l'utilisation accrue d'ordinateurs est associée à une évolution de la demande en faveur de la main-d'œuvre qualifiée (progrès technologique biaisé), des investissements plus importants en formation continue, etc.

entreprises. De plus, nous étudions comment les transformations organisationnelles affectent, à leur tour, l'emploi et les salaires.

Afin d'apporter des réponses à ces questions nous adoptons la forme organisationnelle développée par Garicano (2000). Autrement dit, dans notre modèle, il est optimal d'organiser la production au sein de hiérarchies basées sur les compétences. Néanmoins, nous nous différencions en introduisant des divergences d'intérêts. En effet, nous considérons que les compétences sont déjà acquises par les agents et sont leur information privée. Ainsi, les employés peuvent adopter un comportement stratégique. Un employé peut chercher à travailler le moins possible et/ou cacher la vraie valeur de sa compétence. Dès lors, le principal doit l'inciter à utiliser sa compétence de la façon appropriée. Du point de vue de l'organisation, le coût des compétences est donc associé aux incitations à fournir aux employés. Dans la mesure où ces dernières dépendent des conditions sur le marché du travail, nous faisons le lien entre l'environnement de l'entreprise et sa structure.

Dans notre modèle, nous considérons que les propriétaires d'entreprise embauchent des agents sur le marché du travail, les affectent aux différents postes disponibles dans l'organisation, choisissent l'allocation des problèmes et le montant des salaires. Les agents peuvent être issus de deux distributions, celle des diplômés et celle des non-diplômés. Au sein de chaque population les agents sont hétérogènes et la productivité exacte de chacun est son information privée. Néanmoins, un agent diplômé est plus compétent que n'importe lequel des agents non-diplômés. Ainsi, dans une hiérarchie basée sur les compétences avec trois niveaux nous avons des travailleurs non-diplômés, des cadres diplômés et le propriétaire de l'entreprise capable de réussir les problèmes les plus difficiles. Nous combinons des aspects d'aléa moral du côté des employés et de sélection adverse concernant les caractéristiques à la fois des employés et des chômeurs. Nous restreignons l'ensemble des contrats à disposition du propriétaire en supposant qu'il ne peut proposer qu'un contrat par niveau hiérarchique. Afin de motiver les employés à travailler, l'employeur propose un salaire fixe payé quelle que soit la performance de l'agent et un seuil de performance minimale à atteindre afin de garder son poste dans l'entreprise. Ainsi, lorsque le taux de chômage est faible le licenciement n'est pas une punition forte car il est relativement facile de retrouver du travail. La conséquence en est l'accroissement du salaire. De plus, le taux de chômage affecte la productivité espérée des agents. Au fil du temps, l'employeur licencie les employés (moins compétents à l'équilibre) dont la performance est en dessous du seuil imposé. Ainsi, à l'équilibre stationnaire du marché du travail chaque distribution initiale des productivités est divisée en deux distributions endogènes : la première est celle des agents au sein de l'entreprise et la seconde, celle des chômeurs. Si le taux de chômage est faible, les employés sont moins sélectionnés (la distribution de leurs compétences est plus proche de la distribution initiale) et donc moins productifs.

Les conditions sur le marché du travail affectent à la fois le coût et la productivité de la main-d'œuvre. Or, les décisions d'embauche et d'allocation des problèmes sont déterminées par les coûts relatifs et les productivités relatives des compétences des agents des deux populations. Ainsi, la structure de l'organisation s'adapte aux évolutions de son environnement.

Ce chapitre comporte 7 sections. Dans les Sections 2 et 3, nous présentons respectivement le cadre du modèle et le lien entre les variables de choix du principal : allocation des problèmes, nombre d'employés et salaires. La Section 4 est consacrée à l'analyse du programme d'un principal individuel, alors que la Section 5 caractérise l'équilibre sur le marché du travail. Dans la Section 6, nous présentons nos principaux résultats. La Section 7 conclut. Toutes les preuves sont en Annexe.

2 Le cadre d'analyse

A la suite de Garicano (2000), nous représentons l'entreprise comme une organisation qui traite des problèmes de difficultés x variées. La productivité

θ d'un agent correspond à sa capacité à traiter ces problèmes. L'entreprise s'organise en plusieurs niveaux hiérarchiques ; les niveaux plus élevés sont constitués d'agents plus productifs, capables de résoudre des problèmes de difficulté supérieure. Ainsi, les agents peu productifs doivent transmettre certains des problèmes qui leur parviennent à leurs supérieurs plus productifs. Le niveau le plus élevé de l'organisation est occupé par le propriétaire de l'entreprise, le plus productif de tous.

A la différence de Garicano, nous supposons que la productivité de chaque agent θ est son information privée ; les autres agents ne peuvent observer que son diplôme, qui en est un signal imparfait. En conséquence, les travailleurs adoptent un comportement stratégique. Un travailleur peut chercher à travailler le moins possible ou à cacher à ses supérieurs la vraie valeur de sa productivité. Les contrats en vigueur dans l'entreprise doivent donc respecter de nouvelles contraintes d'incitation.

La production. La production de l'entreprise correspond au nombre de problèmes résolus. Tous les problèmes résolus ont la même valeur, normalisée à 1. La difficulté x d'un problème est inconnue *ex ante*. Le traitement du problème nécessite du temps, son succès nécessite de la compétence et de l'effort. Chaque membre de l'organisation dispose d'une unité de temps par période et nous supposons que :

Hypothèse 1.1. *Un agent peut traiter un problème par unité de temps.*

Un employé de type θ peut résoudre un problème en effectuant un effort uniquement s'il en est capable, c'est-à-dire si la difficulté du problème ne dépasse pas sa productivité ($x \leq \theta$).

Hypothèse 1.2. *Un employé qui reçoit un problème de difficulté supérieure à sa productivité ($x > \theta$) n'est en mesure ni de le résoudre, ni même d'identifier sa difficulté exacte.*

La fonction du coût de l'effort pour un agent θ s'écrit[7] :

$$c(x, \theta) = \begin{cases} c & si \quad x \leq \theta \\ +\infty & si \quad x > \theta \end{cases}$$

A ce stade, nous procédons à deux normalisations. Premièrement, nous fixons le coût de l'effort du propriétaire à zéro. Deuxièmement, la difficulté des problèmes est distribuée selon une loi uniforme, sur le support $[0, 1]$.

Au sein de l'entreprise la communication est possible. Ainsi, un employé incapable de résoudre un problème peut le transmettre (demander de l'aide) à un de ses collègues. Sous l'Hypothèse 1.2, l'employé ne sait pas si son collègue peut résoudre le problème en question. Or, pour traiter un problème transmis, la personne l'ayant reçu doit y consacrer toute son unité de temps, et ce indépendamment du fait qu'à son tour elle soit capable ou non de le résoudre.

Sous les Hypothèses 1.1 et 1.2, il est donc coûteux (en termes de temps) de trouver la personne capable de résoudre chacun des problèmes arrivés dans l'organisation. En conséquence, la forme organisationnelle optimale, comme le démontre Garicano (2000), est une hiérarchie basée sur les compétences (*knowledge based hierarchy*). Dans une telle hiérarchie, les employés les moins compétents occupent le niveau le plus bas de l'organisation. Ils reçoivent les problèmes et en résolvent les plus simples. Les autres sont transmis au niveau suivant où les employés sont capables de résoudre des problèmes plus difficiles. Ainsi, chaque problème suit un processus de transmission du niveau le plus bas à travers les niveaux supérieurs, jusqu'à ce qu'il soit résolu.

---

[7] Les implications de cette forme spécifique de la fonction de coût sont discutées dans la Section 3.1 et en Appendice 9.1.

L'économie. Nous considérons qu'il existe un continuum de masse un de propriétaires d'entreprise[8], que nous appelons également principaux. Ils sont les seuls à avoir accès à la technologie de production. Chacun d'entre eux a la possibilité de produire seul ou d'embaucher des agents dans le continuum d'employés potentiels. Ces derniers peuvent produire uniquement s'ils sont engagés par un propriétaire d'entreprise et ne peuvent en aucun cas travailler à leur compte. Les propriétaires d'entreprise sont homogènes et leur productivité est notée par θ^P. Elle est supérieure à celle de n'importe lequel des employés potentiels. Ceux-ci peuvent être identifiés en fonction de leur niveau de diplôme.

Hypothèse 1.3. *Il y a deux niveaux de diplômes que nous appelons $D = T, C$.*

Nous considérons que les agents $D = C$ sont diplômés, alors que les agents $D = T$ sont non-diplômés. N_D désigne la taille de la population des agents de diplome D. Le diplôme est un signal imparfait de la productivité θ d'un agent particulier et la distribution des productivités à chaque niveau de diplôme est connue de tous. Les θ des non-diplômés sont distribués sur le support $[\underline{\theta}_T, \overline{\theta}_T]$ et leur fonction de répartition est notée $U(\theta)$. Les diplômés, quant à eux, sont distribués sur le support $[\underline{\theta}_C, \overline{\theta}_C]$ et leur fonction de répartition est $S(\theta)$. Afin de simplifier l'analyse nous faisons l'hypothèse suivante :

Hypothèse 1.4. *Les supports des distributions de productivités sont disjoints :*

$$\underline{\theta}_C > \overline{\theta}_T$$

Un diplômé est donc toujours plus productif que n'importe lequel des non-diplômés.

[8] La taille de chaque population est donnée de façon exogène. Les possibilités d'entrée de nouvelles entreprises ou de sortie de celles présentes sur le marché ne sont pas considérées. Cependant, nous discutons les implications potentielles d'une hypothèse de libre entrée lorsque nous présentons les résultats du modèle.

Enfin, nous supposons que tous les agents de l'économie vivent infiniment et ont le même taux d'escompte, noté δ.

Dans ce chapitre, nous considérons que les organisations sont des hiérarchies de trois niveaux. Chaque entreprise est donc constituée de travailleurs, de cadres et du propriétaire. Le choix optimal du nombre de niveaux hiérarchiques est analysé dans le Chapitre 2 de cette thèse[9].

Dans une hiérarchie à trois niveaux, tous les travailleurs sont des agents de diplôme $D = T$ et tous les cadres sont des agents $D = C$. En effet, puisque les non-diplômés sont moins productifs que les diplômés, les travailleurs sont tous embauchés dans la population T et les cadres dans la population C. Sous l'Hypothèse 1.4, l'employeur n'a jamais intérêt à promouvoir un non-diplômé. De plus, dans la Section 3.1, nous montrons qu'il n'est pas dans son intérêt de rétrograder un diplômé.

3 Hiérarchie basée sur les compétences

Le propriétaire de l'entreprise embauche des agents sur le marché du travail et les affecte aux postes disponibles. Les personnes non embauchées restent au chômage. Nous supposons que l'employeur offre le même contrat à tous les agents affectés à un même niveau hiérarchique. Afin de les inciter à travailler, il propose un salaire fixe payé quelle que soit la performance de l'agent, accompagné d'un seuil d'exigence que l'employé doit atteindre afin de garder son poste[10]. Dans la mesure où les agents sont hétérogènes, certains parmi eux ne sont pas capables d'atteindre ce seuil. Ceux-là sont licenciés et remplacés par des chômeurs. Ainsi, les agents qui sont dans l'entreprise sont sélectionnés et la distribution de leur productivité dépend de la distribution des chômeurs et des seuils d'exigence fixés par le principal. Enfin, le

[9] Dans un cadre simplifié où les agents au sein de chacune des trois populations sont homogènes, nous étudions l'impact des conditions sur le marché du travail sur le nombre de niveaux hiérarchiques adoptés par les entreprises.

[10] Ce contrat est similaire à celui proposé par Shapiro et Stiglitz (1984) dans leur analyse du salaire d'efficience.

nombre d'agents que l'employeur embauche à chaque niveau de la hiérarchie dépend, à son tour, de la productivité espérée des employés.

Dans la suite de cette section, nous procédons en trois étapes. Tout d'abord, dans 3.1, nous présentons les schémas d'incitation utilisés par le principal et les problèmes résolus par les employés. Ensuite, dans la Sous-section 3.2, nous caractérisons les distributions stationnaires des employés (travailleurs et cadres). Enfin, dans 3.3, nous présentons le lien entre les seuils de performance et le nombre d'agents embauchés à chaque niveau de la hiérarchie. Nous y obtenons également la production réalisée par l'entreprise.

3.1 Incitations et problèmes résolus

Au début de chaque période, un propriétaire ayant des postes vacants à pourvoir et des chômeurs se rencontrent. Le principal propose des contrats aux agents. Chacun d'entre eux accepte ou refuse. Si un agent refuse le contrat, le poste reste vacant pour la période ; l'agent, quant à lui, retourne dans la population des chômeurs, où il reçoit une indemnité de chômage z_i exogène[11], contingente au dernier poste qu'il a occupé $i = \{T, C\}$. Si l'agent accepte le contrat, il reçoit un problème et décide de le résoudre (exercer l'effort) ou de le transmettre. Son utilité à une période donnée lorsqu'il effectue l'effort s'écrit $(s - c)$, où s est sa rémunération et c le coût de l'effort. A la fin de la période (une fois que l'ensemble des problèmes a été traité), le principal observe la difficulté de chaque problème, l'agent qui l'a reçu et ce qu'il en a fait. Le contrat est alors mis en œuvre. Avec probabilité $(1 - \alpha)$ l'agent quitte l'entreprise pour des raisons exogènes.

Hypothèse 1.5. *Le taux de rotation exogène $(1 - \alpha)$ est indépendant de la productivité des agents et de leur position dans la hiérarchie.*

[11] Nos résultats ne sont pas qualitativement affectés si l'indemnité de chômage est contingente au salaire perçu par l'agent dans son dernier emploi.

Nous imposons deux restrictions à l'ensemble des contrats proposés par le principal. Premièrement, il y a un seul contrat par niveau hiérarchique[12], c'est-à-dire qu'il y a un contrat pour les travailleurs et un contrat pour les cadres. Deuxièmement, les incitations à l'effort sont garanties par un salaire fixe et la menace de licencier toute personne dont la performance est en dessous d'un certain seuil. Dans ce chapitre, la forme des contrats est fixée. Dans le Chapitre 3, nous traitons la question du contrat optimal lorsque les agents sont hétérogènes, la productivité de chacun est son information privée et sa performance est observable pour les parties contractantes mais ne peut pas être vérifiée par un tiers. Nous y discutons les raisons pour lesquelles proposer un menu de contrats aux agents peut être sous-optimal pour le principal. Concernant l'utilisation du salaire fixe et la menace de licenciement, nous montrons qu'en présence de chômage, ce type d'incitations est optimal au moins pour une partie des problèmes à effectuer. Dans la mesure où les résultats de ce chapitre reposent sur le lien entre les conditions sur le marché du travail et le coût des salaires, même si le licenciement n'est pas le seul moyen d'incitation utilisé par le principal, nous pensons que nos principales conclusions devraient rester valables.

Ainsi, le contrat proposé aux agents du niveau $i = \{T, C\}$ se présente sous la forme (s_i, x_i). L'employé reçoit le salaire fixe s_i. A la fin de la période il est licencié si sa performance est inférieure au seuil d'exigence x_i, c'est-à-dire s'il n'a pas résolu un problème de difficulté $x \leq x_i$. A chaque période, une partie des travailleurs dont la productivité est en-dessous du seuil d'exigence est détectée par l'employeur et licenciée. Ainsi, un employé peut se retrouver au chômage pour trois raisons : séparation exogène, ou licenciement, soit parce qu'il n'a pas travaillé, soit parce que sa faible compétence a été détectée. Nous supposons que le temps passé par un agent au chômage et les raisons pour lesquelles il est chômeur ne sont pas observées.

[12] Sous cette hypothèse, la question de l'apprentissage dynamique du principal ne se pose pas. En effet, le contrat unique doit être optimal compte tenu de la distribution des productivités des employés (travailleurs ou cadres).

Hypothèse 1.6. *Le marché du travail est anonyme.*

Le diplôme est la seule information dont disposent les employeurs concernant les chômeurs. Un agent au chômage reçoit une indemnité et à la fin de chaque période avec une certaine probabilité retrouve du travail. Sous l'Hypothèse 1.6, cette probabilité dépend uniquement du fait que l'agent soit diplômé ou non et non pas de ses expériences passées. Elle est donc indépendante de θ. La probabilité de réembauche d'un chômeur non-diplômé est notée λ_T, celle d'un chômeur diplômé : λ_C. Ces probabilités sont endogènes et sont déterminées dans la Section 5.

Donc, si $V(\theta, i)$ désigne l'utilité inter-temporelle espérée d'un agent de type θ embauché au niveau i, nous pouvons écrire l'utilité inter-temporelle espérée d'un chômeur non-diplômé de type θ ($V_U(\theta, T)$) :

$$V_U(\theta, T) = z_T + \delta[\lambda_T V(\theta, T) + (1 - \lambda_T)V_U(\theta, T)] \tag{1.1}$$

Celle d'un chômeur diplômé de type θ ($V_U(\theta, C)$) s'écrit :

$$V_U(\theta, C) = z_C + \delta[\lambda_C V(\theta, C) + (1 - \lambda_C)V_U(\theta, C)] \tag{1.2}$$

Afin de simplifier l'exposé des résultats, nous adoptons la notation suivante : soit $G(t)$ la fonction de répartition d'une variable aléatoire t distribuée sur le support $[\underline{t}, \overline{t}]$, alors $\bar{G}(y) = \int_{\underline{t}}^{y} G(t)dt$.

Lemme 1.1. $\bar{G}(y) = \int_{\underline{t}}^{y} (y - t)g(t)dt$

La preuve du lemme est directe à partir d'une intégration par parties.

3.1.1 Les travailleurs.

Rémunération

Considérons, tout d'abord, le cas d'un travailleur dont la productivité est inférieure au seuil d'exigence fixé par le principal ($\theta < x_T$). A chaque période, il se trouve dans une des situations suivantes :

- Il reçoit un problème $x \leq \theta$. S'il effectue l'effort, il le résout et garde son poste avec probabilité α. S'il choisit de ne pas travailler, il est licencié à la fin de la période. En supposant que le principal l'incite à traiter tout problème $\theta < x_T$, la probabilité que l'agent résolve un problème et dépense donc le coût de l'effort est $Prob(x \leq \theta) = \theta$.
- Il reçoit un problème $x > \theta$ et le transmet au niveau hiérarchique suivant :
  - si $x \in [\theta, x_T]$ l'agent est licencié pour cause de performance insuffisante,
  - si $x > x_T$ il garde son poste avec probabilité α.

Lorsqu'un agent reçoit un problème $x > \theta$, d'après l'Hypothèse 1.2 il n'en connaît pas la difficulté exacte. Dès lors, il est dans son intérêt de le transmettre. En effet, s'il ne le transmet pas il est sûr d'être licencié. En revanche, en cas de transmission, avec une probabilité strictement positive le problème est $x > x_T$, auquel cas l'agent garde son poste avec probabilité α.

Ainsi, l'utilité inter-temporelle espérée d'un travailleur $\theta < x_T$ qui effectue l'effort s'écrit :

$$V(\theta,T) = s_T - \theta c + \delta(1-\alpha)V_U(\theta,T) + \alpha\delta[(1-x_T+\theta)V(\theta,T) + (x_T-\theta)V_U(\theta,T)] \quad (1.3)$$

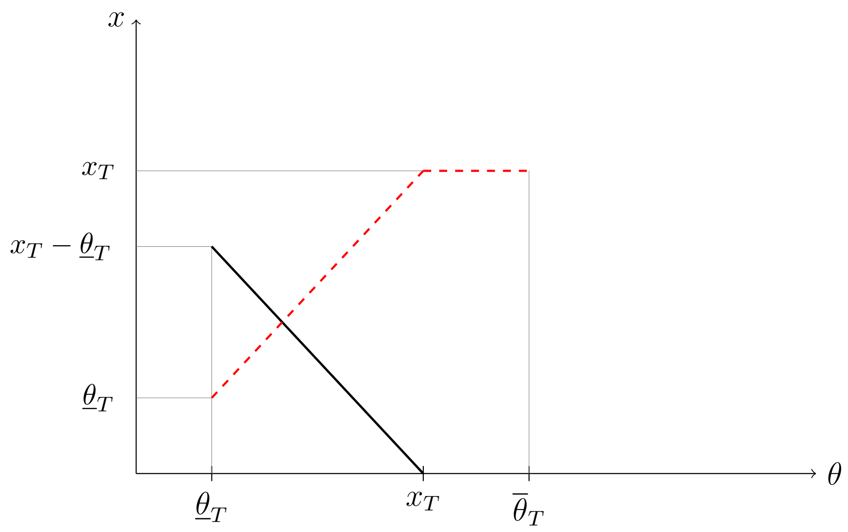
A présent nous nous intéressons à un travailleur dont la productivité est supérieure au seuil d'exigence, $\theta \in [x_T, \overline{\theta}_T]$. A chaque période, il se trouve dans une des situations suivantes :

- Il reçoit un problème $x \leq x_T$. S'il travaille, il garde son poste avec probabilité α, dans le cas contraire il est licencié à la fin de la période. Par conséquent la probabilité qu'il dépense le coût de l'effort est x_T.
- Il reçoit un problème $x > x_T$, il le transmet et garde son poste avec probabilité α.

Ainsi, l'utilité inter-temporelle espérée d'un travailleur $\theta \geq x_T$ qui effectue l'effort s'écrit :

$$V(\theta, T) = s_T - x_T c + \delta\alpha V(\theta, T) + \delta(1-\alpha)V_U(\theta, T) \qquad (1.4)$$

Sur la Figure 1.1, nous récapitulons les probabilités pour un travailleur d'être licencié ou de dépenser le coût de l'effort en fonction de sa productivité θ.



— Probabilité pour un travailleur θ d'être licencié.
- - Probabilité pour un travailleur θ de dépenser le coût de l'effort.

Figure 1.1 : La probabilité pour un travailleur θ d'être licencié ou de dépenser le coût de l'effort.

Lorsqu'un travailleur reçoit un problème qu'il peut et est censé résoudre (c'est-à-dire $x \leq \min\{\theta, x_T\}$), il effectue l'effort si et seulement si :

$$s_T - c + \delta\alpha V(\theta, T) + \delta(1-\alpha)V_U(\theta, T) \geq s_T + \delta V_U(\theta, T)$$

$$\Leftrightarrow V(\theta, T) - V_U(\theta, T) \geq \frac{c}{\alpha\delta} \qquad (1.5)$$

Comme chez Shapiro et Stiglitz (1984), un agent est incité à travailler lorsque la rente associée au fait de garder son travail est suffisamment élevée.

Problèmes résolus

Comme nous venons de le voir, sous la contrainte d'incitation (1.5), un travailleur résout un problème s'il le peut ($x \leq \theta$) et si le principal lui a imposé de le faire ($x \leq x_T$). La fonction de répartition des productivités des travailleurs est notée $F(\theta)$, sa détermination endogène est discutée dans la Section 3.2. Ainsi, la fraction de problèmes résolus par les travailleurs s'écrit :

$$\begin{aligned} Prob(x \leq \min\{\theta, x_T\}) &= \int_{\underline{\theta}_T}^{x_T} \theta f(\theta) d\theta + \int_{x_T}^{\bar{\theta}_T} x_T f(\theta) d\theta \\ &= \int_{\underline{\theta}_T}^{\bar{\theta}_T} x_T f(\theta) d\theta - \int_{\underline{\theta}_T}^{x_T} (x_T - \theta) f(\theta) d\theta \end{aligned}$$

Ce qui, en utilisant le Lemme 1.1, nous donne :

$$x_T - \bar{F}(x_T)$$

Du fait que les agents sont hétérogènes dans leur productivité, une partie d'entre eux n'arrivent pas à atteindre le seuil d'exigence x_T. Il s'agit des travailleurs ayant reçu un problème de difficulté comprise entre θ et x_T. Ces employés sont licenciés par le principal. La proportion de travailleurs licenciés à la fin de chaque période est donnée par :

$$l_T = \bar{F}(x_T) \tag{1.6}$$

Les problèmes non résolus par les travailleurs sont transmis au niveau hiérarchique supérieur, celui des cadres. La fraction de problèmes reçus par les cadres s'écrit alors :

$$r_C = 1 - x_T + l_T \tag{1.7}$$

Cette fraction dépend du seuil d'exigence des travailleurs et des caractéristiques de leur distribution.

3.1.2 Les cadres.

Rémunération

Afin de définir l'utilité inter-temporelle d'un cadre nous devons préciser la probabilité pour lui de dépenser le coût de l'effort et sa punition dans le cas de performance insuffisante.

Les cadres se spécialisent dans le traitement des problèmes que les travailleurs n'ont pas su résoudre. S'il y a n travailleurs par cadre, alors en moyenne un cadre reçoit nr_C problèmes. Dans la mesure où il ne peut traiter qu'un problème par unité de temps, le nombre de problèmes traités est $\min\{nr_C, 1\}$.

La probabilité que la difficulté d'un problème parmi ceux transmis par les travailleurs soit inférieure à X, si $X < x_T$ est :

$$Z(X) = \frac{\bar{F}(X)}{r_C}$$

Dans le cas où $X > x_T$, cette probabilité s'écrit :

$$Z(X) = \frac{X - x_T + \bar{F}(x_T)}{r_C}$$

Sous l'Hypothèse 1.4, la productivité de tout agent issu de la population $D = C$ est supérieure à x_T. Donc, la probabilité qu'un des problèmes transmis soit de difficulté inférieure à la productivité θ d'un cadre est :

$$Z(\theta) = \frac{\theta - x_T + \bar{F}(x_T)}{r_C} \tag{1.8}$$

Un cadre dépense le coût de l'effort avec probabilité :

$$\min\{nr_C, 1\} Z(\min\{\theta, x_C\}) \tag{1.9}$$

Notons que cette probabilité dépend de la productivité des travailleurs. Lorsque ceux-ci sont plus productifs, c'est-à-dire r_c est faible, la probabilité pour un cadre de dépenser le coût de l'effort est plus faible. Par conséquent, le salaire des cadres va dépendre de leur propre seuil d'exigence, mais également de celui que le principal choisit pour les travailleurs.

A présent, nous caractérisons la punition d'un cadre ayant transmis un problème $x \leq x_C$. Le principal a la possibilité de le licencier ou de le rétrograder au poste de travailleur.

Lemme 1.2. *Un cadre qui transmet un problème $x \leq x_C$ est licencié par l'employeur.*

Un cadre dont la performance n'est pas satisfaisante accepte d'être rétrogradé seulement si son utilité en tant que travailleur excède celle qu'il peut espérer en tant que chômeur. Or, dans ce cas de figure, le coût incitatif pour le principal est plus faible s'il licencie l'agent. Si en revanche, être rétrogradé est une punition plus forte et donc préférable pour le principal, le cadre préfère quitter l'entreprise plutôt que d'accepter le nouveau poste. Dès lors, la punition maximale que l'employeur est en mesure d'infliger à l'agent est le licenciement.

Sur la Figure 1.2 nous représentons la probabilité pour un agent θ (travailleur ou cadre) de dépenser le coût de l'effort ou d'être licencié.

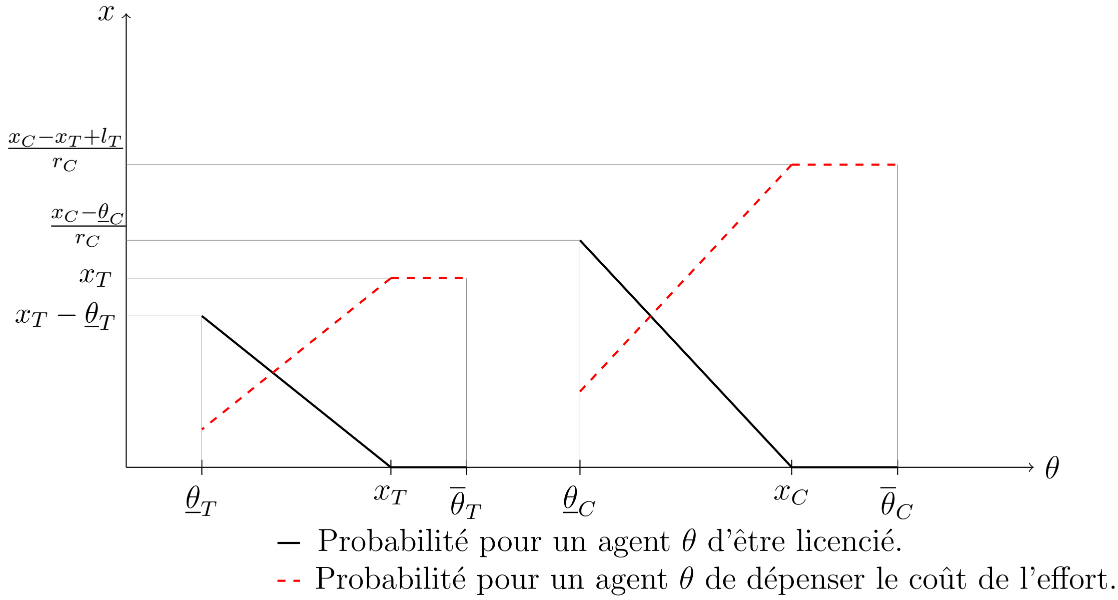


Figure 1.2 : La probabilité pour un agent θ d'être licencié ou de dépenser le coût de l'effort dans le cas où $n = \frac{1}{r_C}$.

Nous pouvons écrire à présent l'utilité inter-temporelle espérée d'un cadre dont la productivité est inférieure au seuil d'exigence ($\theta < x_C$) qui effectue l'effort

$$\begin{aligned} V(\theta, C) = s_C - \min\{nr_C, 1\}Z(\theta)c \quad &+\delta(1-\alpha)V_U(\theta, C) \\ &+\delta\alpha[(1 - \min\{nr_C, 1\}(Z(x_C) - Z(\theta)))V(\theta, C) \\ &+ \min\{nr_C, 1\}(Z(x_C) - Z(\theta))V_U(\theta, C)] \end{aligned} \tag{1.10}$$

et celle d'un cadre plus productif que le seuil x_C :

$$V(\theta, C) = s_C - \min\{nr_C, 1\}Z(x_C)c + \delta(1-\alpha)V_U(\theta, C) + \alpha\delta V(\theta, C) \tag{1.11}$$

Comme pour les travailleurs, lorsqu'un cadre reçoit un problème qu'il peut et est censé résoudre (c'est-à-dire $x \leq \min\{\theta, x_C\}$), il effectue l'effort si et seulement si la rente associée au fait de garder son travail est suffisamment élevée :

$$s_C - c + \delta\alpha V(\theta, C) + \delta(1-\alpha)V_U(\theta, C) \geq s_C + \delta V_U(\theta, C)$$

$$V(\theta, C) - V_U(\theta, C) \geq \frac{c}{\alpha\delta} \tag{1.12}$$

Dans la mesure où le coût de l'effort est indépendant de θ et où la probabilité de retrouver du travail est la même pour tout chômeur, à niveau de diplôme donné, notre modèle présente la propriété suivante.

Lemme 1.3. *Les salaires incitatifs ne dépendent pas de θ.*

La preuve formelle est exposée en Annexe 9.1. Le principal sature les contraintes d'incitation (1.5) et (1.12). La probabilité de réembauche étant indépendante de θ, il est facile de montrer que l'utilité inter-temporelle d'un chômeur[13] ne dépend pas de θ. Concernant $V(\theta, D)$, un θ plus élevé (lorsque $\theta < x_i$) augmente d'un côté la probabilité que l'agent garde son poste (le bénéfice associé est $\alpha\delta(V(\theta, D) - V_U(\theta, D)))$ et de l'autre accroît la probabilité que le coût de l'effort c soit dépensé. Lorsque la contrainte d'incitation est saturée, le coût marginal est égal à la recette marginale et l'utilité inter-temporelle du travailleur ne dépend pas de θ.

Compte tenu de ces résultats, nous simplifions les notations comme suit pour le reste du chapitre : $V(T)$ (à la place de $V(\theta, T)$), $V_U(T)$ (à la place de $V_U(\theta, T)$), $V(C)$ (à la place de $V(\theta, C)$), $V_U(C)$ (à la place de $V_U(\theta, C)$).

Problèmes résolus

Lorsque le salaire des cadres vérifie la contrainte d'incitation (1.12), un cadre résout un problème qu'il a reçu s'il peut et est censé le faire, c'est-à-dire si $x \leq \min\{\theta, x_C\}$. La fonction de répartition des productivités des cadres est notée $\Phi(\theta)$ ($\varphi(\theta)$ est la fonction de distribution correspondante). Comme pour celle des travailleurs, sa

[13] Prenons le cas des non-diplômés $V_U(\theta, D) = z_D + \delta V_U(\theta, D) + \lambda_D\delta(V(\theta, D) - V_U(\theta, D)) \Leftrightarrow V_U(\theta, D)(1-\delta) = z_D + \frac{\lambda_D c}{\alpha}$.

détermination endogène est discutée dans la Section 3.2. La fraction de problèmes que les cadres arrivent à résoudre parmi ceux qu'ils ont reçu est donnée par :

$$\int_{\underline{\theta}_C}^{x_C} Z(\theta)\varphi(\theta)d\theta + \int_{x_C}^{\overline{\theta}_C} Z(x_C)\varphi(\theta)d\theta = Z(x_C) - \frac{\bar{\Phi}(x_C)}{r_C}$$

Les cadres ayant reçu des problèmes de difficulté $x \in]\theta, x_C]$ ne sont pas en mesure de satisfaire l'exigence du principal et sont licenciés. La proportion de cadres licenciés à chaque période est donc :

$$l_C = \frac{\bar{\Phi}(x_C)}{r_C}\min\{nr_C, 1\} \tag{1.13}$$

Nous pouvons à présent caractériser la fraction de problèmes reçus par le principal. Un problème qui arrive dans l'entreprise (donc tiré de la distribution uniforme sur $[0,1]$) est transmis au principal si ni les travailleurs ni les cadres n'ont été en mesure de le résoudre. Sous l'Hypothèse 1.4, un problème non résolu par un cadre ne peut pas l'être par un travailleur. Par conséquent, la proportion de problèmes reçus par le propriétaire s'écrit :

$$\begin{aligned} r_P &= 1 - \int_{\underline{\theta}_C}^{x_C} \theta\varphi(\theta)d\theta - \int_{x_C}^{\overline{\theta}_C} x_C\varphi(\theta)d\theta \\ &= 1 - x_C + \bar{\Phi}(x_C) \end{aligned} \tag{1.14}$$

3.2 Sélection

Caractérisons maintenant les distributions stationnaires des productivités des travailleurs et des cadres.

Le choix des seuils de performance affecte les caractéristiques des distributions des productivités des employés. En effet, x_T et x_C déterminent la proportion d'agents licenciés et remplacés à la fin de chaque période.

A l'état stationnaire, la distribution des travailleurs à la fin de la période, une fois que les personnes licenciées ou parties pour des raisons exogènes ont été remplacées par des chômeurs non-diplômés, doit être la même qu'au début de la période. Ainsi, à distribution des chômeurs non-diplômés $Q(\theta)$ donnée, nous obtenons la condition suivante qui garantit la stationnarité de $F(\theta)$.

$$Q(\theta) = \begin{cases} \dfrac{(1-\alpha)F(\theta) + \alpha \displaystyle\int_{\underline{\theta}_T}^{\theta} (x_T - u) f(u) du}{(1-\alpha+\alpha l_T)} & \text{si} \quad \theta \in [\underline{\theta}_T, x_T] \\ \dfrac{(1-\alpha)F(\theta) + \alpha l_T}{(1-\alpha+\alpha l_T)} & \text{si} \quad \theta \in [x_T, \overline{\theta}_T] \end{cases} \tag{1.15}$$

Les détails du calcul se trouvent dans l'Annexe 9.2.

Le même raisonnement que dans le cas des travailleurs s'applique aux cadres. Ainsi, le lien à l'état stationnaire entre la distribution des cadres ($\Phi(\theta)$) et celle des chômeurs diplômés ($P(\theta)$) est donné par :

$$P(\theta) = \begin{cases} \dfrac{(1-\alpha)\Phi(\theta) + \alpha \displaystyle\int_{\underline{\theta}_C}^{\theta} \frac{(x_C - u)\varphi(u)}{r_C} du}{(1-\alpha+\alpha l_C)} & \text{si} \quad \theta \in [\underline{\theta}_C, x_C] \\ \dfrac{(1-\alpha)\Phi(\theta) + \alpha l_C}{(1-\alpha+\alpha l_C)} & \text{si} \quad \theta \in [x_C, \overline{\theta}_C] \end{cases} \tag{1.16}$$

Notons que la distribution des chômeurs non-diplômés dépend uniquement du seuil d'exigence des travailleurs. En revanche, la distribution des chômeurs diplômés dépend à la fois de x_C et de x_T. Effectivement, le choix de x_T affecte la distribution des problèmes reçus par les cadres et par conséquent la probabilité de chacun d'être licencié.

Lemme 1.4. *La distribution des productivités des employés θ domine, au sens de la dominance stochastique d'ordre un, la distribution correspondante des chômeurs :*

$$F(\theta) \succ_1 Q(\theta) \qquad et \qquad \Phi(\theta) \succ_1 P(\theta)$$

En effet, à l'état stationnaire la distribution des chômeurs est constituée d'employés partis pour des raisons exogènes et d'employés licenciés à cause de leur performance insuffisante.

3.3 Production et nombre d'employés

La production étant organisée dans des hiérarchies basées sur les compétences, les problèmes arrivent dans l'entreprise au niveau des travailleurs. Les cadres et le principal sont chargés de traiter les problèmes non résolus par leurs subalternes. Chaque travailleur peut traiter un seul problème par période, donc le nombre de problèmes qui arrivent dans l'entreprise correspond au nombre de travailleurs n_T. Ces derniers en résolvent une partie $(1 - r_C)$ et en transmettent l'autre (r_C) aux cadres. Ainsi, $n_T r_C$ problèmes arrivent au deuxième niveau de la hiérarchie. Chaque cadre peut traiter un problème transmis par période. Il est sous optimal pour le principal d'engager des cadres qui n'utilisent pas pleinement leur temps. Dès lors, le nombre de cadres (n_C) embauchés est obtenu en saturant leur contrainte de temps :

$$n_T r_C = n_C \tag{1.17}$$

n_C dépend à la fois du nombre de travailleurs et de leur performance. En effet, des travailleurs plus autonomes (seuil d'exigence x_T plus élevé et/ou « meilleure » distribution des productivités, c'est-à-dire faible l_T) résolvent plus de problèmes et par conséquent moins de cadres sont nécessaires pour les aider. Nous remarquons que le nombre de travailleurs par cadre est $n = \dfrac{n_T}{n_C} = \dfrac{1}{r_C}$.

La taille de l'entreprise est limitée par la présence d'un seul principal doté d'une unité de temps et par le degré de traitement que chaque problème arrivé dans l'entreprise doit recevoir. Ici, nous supposons que chaque problème doit recevoir un traitement maximal. Cependant, nos résultats ne seraient pas qualitativement affectés par l'introduction d'un coût associé au fait qu'un problème ne soit pas traité et un choix endogène de la proportion de problèmes à résoudre. Parmi les n_T

problèmes reçus par l'entreprise, la proportion qui n'est résolue ni par les travailleurs ni par les cadres est r_P. A partir de la contrainte de temps du principal, nous avons :

$$n_T r_P = 1 \tag{1.18}$$

Ainsi, le nombre de travailleurs, que le principal décide d'engager, dépend de la probabilité qu'un problème soit transmis au principal.

Sur la Figure 1.3 nous représentons les flux de problèmes à travers la hiérarchie.

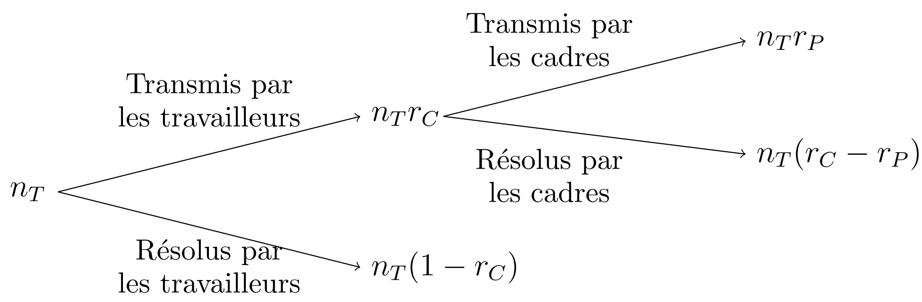


Figure 1.3 : Le nombre de problèmes résolus et transmis à chaque niveau de la hiérarchie

4 Le problème du principal

Chaque propriétaire d'entreprise est « trop petit » pour avoir un impact sur les conditions du marché. Donc, chacun d'entre eux maximise son profit à options de sortie et fonctions de répartition des chômeurs données. Celles-ci sont déterminées de façon adéquate à l'équilibre de marché. Dans cette section, nous présentons les arbitrages pour le principal lorsqu'il est amené à choisir les seuils de performance et discutons l'impact d'évolutions dans l'environnement de l'entreprise sur le choix de ces seuils.

4.1 Le programme

A conditions du marché du travail données, chaque principal choisit les seuils d'exigence, les salaires et le nombre d'agents à embaucher pour chaque poste. Les salaires et le nombre de personnes à chaque niveau hiérarchique sont donnés par les contraintes décrites dans la Section 3.

Tout d'abord, le nombre de travailleurs et de cadres dans l'entreprise en fonction de la politique de licenciement du principal, est obtenu à partir des contraintes de temps. Ainsi, en remplaçant (1.7), (1.13) et (1.14) dans les équations (1.17) et (1.18), nous obtenons :

$$\begin{cases} n_T = \dfrac{1}{1 - x_C + l_C r_C} \\ n_C = \dfrac{1 - x_T + l_T}{1 - x_C + l_C r_C} \end{cases} \tag{1.19}$$

où les proportions d'agents licenciés (l_T et l_C) dépendent des seuils d'exigence correspondants et de la distribution des employés. Dans la Section 3.2, nous avons vu que la distribution stationnaire des travailleurs (*resp.* cadres) est fonction de la distribution stationnaire des chômeurs non-diplômés (*resp.* diplômés). A partir des équations (1.15) et (1.16), nous exprimons les fonctions de distribution respectivement des travailleurs $f(\theta)$ et des cadres $\varphi(\theta)$.

$$\begin{aligned} f(\theta) &= \frac{q(\theta)(1 - \alpha + \alpha l_T)}{1 - \alpha + \alpha(x_T - \theta)} \qquad && \text{pour } \theta < x_T \\ \varphi(\theta) &= \frac{p(\theta)(1 - \alpha + \alpha l_C) r_C}{(1 - \alpha) r_C + \alpha(x_C - \theta)} \qquad && \text{pour } \theta < x_C \end{aligned}$$

Nous pouvons donc réécrire $l_T = \bar{F}(x_T)$ (voir l'équation (1.6)) et $l_C r_C = \bar{\Phi}(x_C)$ (voir l'équation (1.13)) en fonction des distributions des chômeurs respectivement non-diplômés et diplômés :

$$\begin{cases} l_T = \displaystyle\int_{\underline{\theta}_T}^{x_T} \frac{(x_T - \theta) q(\theta)(1 - \alpha + \alpha l_T)}{1 - \alpha + \alpha(x_T - \theta)} d\theta \\ l_C r_C = \displaystyle\int_{\underline{\theta}_C}^{x_C} \frac{(x_C - \theta) p(\theta)((1 - \alpha) r_C + \alpha l_C r_C)}{(1 - \alpha) r_C + \alpha(x_C - \theta)} d\theta \end{cases} \tag{1.20}$$

Enfin, les salaires incitatifs sont obtenus en saturant les contraintes d'incitation (1.5) et (1.12), et en tenant compte, pour celui des cadres, du fait que le nombre de travailleurs par cadre dans l'organisation correspond à $n = \dfrac{n_T}{n_C} = \dfrac{1}{r_C}$.

$$\begin{cases} s_T = \dfrac{c(1-\alpha\delta)}{\alpha\delta} + cx_T + (1-\delta)V_U(T) \\ s_C = \dfrac{c(1-\alpha\delta)}{\alpha\delta} + c\dfrac{(x_C - x_T + l_T)}{r_C} + (1-\delta)V_U(C) \end{cases} \tag{1.21}$$

Donc, le programme du principal consiste à choisir les seuils de performance x_T et x_C de sorte à maximiser son profit :

$$\max_{x_T, x_C} \Pi = n_T(\theta^P - s_T) - s_C n_C$$

où n_T, n_C, s_T, s_C, l_T et $l_C r_C$ sont donnés par les équations (1.19), (1.20) et (1.21).

En remplaçant n_C par $n_T r_C$ (voir équation (1.17)) et en dérivant par rapport à x_C, nous obtenons la condition de premier ordre pour le seuil d'exigence des cadres (CPO_{x_C}) :

$$\frac{\partial \Pi}{\partial x_C} = 0 \quad \Leftrightarrow \frac{\partial n_T}{\partial x_C}(\theta^P - s_T - s_C r_C) - \frac{\partial s_C}{\partial x_C} n_T r_C = 0 \tag{1.22}$$

En effet, $\dfrac{\partial s_T}{\partial x_C} = 0$ (voir (1.21)) et $\dfrac{\partial r_C}{\partial x_C} = 0$ (voir (1.7) et (1.20))

(1.22) se réécrit de la façon suivante :

$$\Big(1 - \frac{\partial (l_C r_C)}{\partial x_C}\Big)\frac{(\theta^P - s_T - s_C r_C)}{r_P} - c = 0 \tag{1.23}$$

Le coût marginal d'un x_C plus élevé est l'accroissement du salaire des cadres. Le gain marginal correspond à l'accroissement du nombre de problèmes qui arrivent dans l'organisation. Résoudre un problème additionnel apporte un bénéfice de $(\theta^P - s_T - s_C r_C)$. Lorsque le principal est plus compétent ou lorsque les salaires des cadres ou des travailleurs sont plus faibles, il devient plus rentable d'accroître le nombre de problèmes traités par l'entreprise.

La condition de premier ordre pour le seuil d'exigence des travailleurs (CPO_{x_T}) est :

$$\frac{\partial \Pi}{\partial x_T} = 0 \quad \Leftrightarrow \frac{\partial n_T}{\partial x_T}(\theta^P - s_T - s_C r_C) - n_T \frac{\partial s_T}{\partial x_T} - \frac{\partial r_C}{\partial x_T} s_C n_T - \frac{\partial s_C}{\partial x_T} n_T r_C = 0$$

$$\Leftrightarrow -\frac{\partial (l_C r_C)}{\partial x_T} \frac{(\theta^P - s_T - s_C r_C)}{r_P} - c + \Big(1 - \frac{\partial l_T}{\partial x_T}\Big)\Big(\frac{c}{\alpha\delta} + (1 - \delta V_U(C))\Big) = 0$$

Ce qui se réécrit :

$$\Big(1 - \frac{\partial l_T}{\partial x_T}\Big) K \frac{(\theta^P - s_T - s_C r_C)}{r_P} - c + \Big(1 - \frac{\partial l_T}{\partial x_T}\Big)(\frac{c}{\alpha\delta} + (1 - \delta V_U(C))) = 0 \quad (1.24)$$

où $K = \frac{1}{r_C} \int_{\underline{\theta}_C}^{x_C} \frac{\varphi(\theta)(x_C - \theta)\alpha(x_C - \theta - (l_C r_C))}{(1-\alpha) r_C + \alpha(x_C - \theta)} d\theta$. Dans l'Annexe 9.4, nous présentons les détails du calcul de la dérivée $\frac{\partial (l_C r_C)}{\partial x_T}$ et nous montrons que $K > 0$.

A x_C donné, un seuil d'exigence plus élevé pour les travailleurs augmente la production de l'entreprise. Ainsi, un x_T plus élevé augmente la productivité espérée des cadres, en diminuant $l_C r_C$. Par ailleurs, déléguer plus de tâches aux travailleurs permet de modifier le coût salarial total. D'un côté, cela augmente le salaire des travailleurs, de l'autre, cela réduit la masse salariale versée aux cadres. En effet, en augmentant x_T le principal réduit à la fois le nombre de cadres et leur salaire. Lorsque les travailleurs sont plus autonomes, l'employeur peut engager moins de cadres pour les aider. De plus, à x_C donné, lorsque les travailleurs résolvent plus de problèmes, cela diminue la probabilité pour un cadre de dépenser le coût de l'effort et réduit donc son salaire incitatif.

Pour le reste du chapitre nous faisons l'hypothèse suivante.

Hypothèse 1.7. *Nous supposons que les distributions initiales sont telles que :*

$$\frac{\partial^2 l_T}{\partial x_T^2}(x_T, q(\theta)) > 0 \qquad \textit{pour tout } x_T$$

$$\frac{\partial^2 (l_C r_C)}{\partial x_C^2}(x_C, p(\theta)) > 0 \quad \textit{pour tout } x_C$$

Dans l'Annexe 9.6, nous montrons qu'une condition suffisante pour que cette hypothèse soit vérifiée est que la distribution initiale des productivités soit proche d'une distribution uniforme et que le support des types ne soit pas trop large.

Le gain lié à l'augmentation de x_T et x_C dépend des distributions des travailleurs et des cadres à travers l'accroissement marginal de la proportion des licenciés l_T et l_C (voir équations (1.23) et (1.24)). Dans la mesure où le choix des seuils d'exigence affecte les caractéristiques des distributions des productivités, nous obtenons le résultat suivant :

Lemme 1.5. *Il existe des valeurs pour x_T et x_C pour lesquelles*

$$\frac{\partial l_T}{\partial x_T} > 1 \qquad et \qquad \frac{\partial (l_C r_C)}{\partial x_C} > 1$$

Ainsi, à partir du Lemme 1.5 et de l'Hypothèse 1.7 il est immédiat d'obtenir le lemme suivant.

Lemme 1.6.

- *Il existe $\tilde{x}_T < \overline{\theta}_T$ tel que $\dfrac{\partial l_T}{\partial x_T}(\tilde{x}_T, q(\theta)) = 1$, si $x_T > \tilde{x}_T$ alors $\dfrac{\partial l_T}{\partial x_T} > 1$.*
- *Il existe $\tilde{x}_C < \overline{\theta}_C$ tel que $\dfrac{\partial (l_C r_C)}{\partial x_C}(\tilde{x}_C, p(\theta)) = 1$, si $x_C > \tilde{x}_C$ alors $\dfrac{\partial (l_C r_C)}{\partial x_C} > 1$.*

Ce lemme implique qu'augmenter les seuils d'exigence au delà de $\tilde{x}_T$ et $\tilde{x}_C$ devient contre-productif. En effet, dans ce cas, en augmentant x_T (*resp.* x_C) le principal accroît la proportion de problèmes transmis r_C (*resp.* r_P). En conséquence, les seuils d'exigence ne sont jamais fixés à leurs niveaux maximaux ($\overline{\theta}_T$ pour x_T et $\overline{\theta}_C$ pour x_C), même lorsque le coût de l'effort est proche de zéro.

4.2 L'environnement de l'entreprise et le choix des seuils d'exigence.

Dans cette section, nous nous concentrons sur les effets de changements dans l'environnement économique de l'entreprise sur le choix du principal des seuils d'exigence. Les effets d'équilibre sur l'organisation, l'emploi et les salaires sont discutés dans la Section 6.

Un changement dans l'environnement de l'entreprise affecte le choix des seuils d'exigence. Par ailleurs, lorsqu'un des seuils change, cela affecte le choix de l'autre. En effet, la dérivée croisée du profit par rapport aux seuils d'exigence pour les valeurs d'équilibre s'écrit : $\frac{\partial^2 \Pi}{\partial x_C \partial x_T} = -\frac{\partial^2 l_C r_C}{\partial x_C \partial x_T}\Pi$. Dans le cas où une augmentation du seuil d'exigence des travailleurs réduit la productivité marginale des cadres (c'est-à-dire $\frac{\partial^2 l_C r_C}{\partial x_C \partial x_T} > 0$), les deux seuils sont des substituts $\left(\frac{\partial^2 \Pi}{\partial x_C \partial x_T} < 0\right)$. En revanche, lorsque l'augmentation du seuil d'exigence des travailleurs augmente la productivité marginale des cadres, les deux seuils sont des compléments $\left(\frac{\partial^2 \Pi}{\partial x_C \partial x_T} > 0\right)$. Ici, les deux situations sont possibles en fonction des valeurs des paramètres.

Ainsi, pour chaque évolution dans l'environnement de l'organisation il y a deux effets sur x_T et x_C, l'un direct et l'autre indirect qui transite par l'interaction des seuils d'exigence. Les résultats annoncés dans la proposition le sont dans le cas où cet effet indirect n'est pas trop fort.

Proposition 1.1. ***Statique comparative sur x_T et x_C.***
Si les effets indirects ne sont pas trop forts, alors :

1. *Les deux seuils d'exigence (x_T et x_C) augmentent lorsque (i) la compétence du principal (θ^P) augmente, (ii) ou l'option de sortie des non-diplômés ($V_U(T)$) diminue.*

2. x_T diminue et x_C augmente lorsque l'option de sortie des diplômés ($V_U(C)$) diminue.

Dans l'Annexe 9.7, nous démontrons cette proposition et précisons les conditions sous lesquelles nos résultats sont obtenus.

Une augmentation de la proportion de problèmes que le principal est capable de résoudre (θ^P) ou une réduction de l'option de sortie des non-diplômés ($V_U(T)$) ont pour effet d'augmenter le gain marginal du principal lorsqu'un problème supplémentaire est résolu. Dès lors, le principal augmente les seuils de performance afin d'accroître le nombre de problèmes traités par l'organisation. Une diminution de l'option de sortie des diplômés augmente le seuil d'exigence des cadres puisque le gain marginal de la résolution d'un problème additionnel est plus élevé. D'un autre côté, un coût plus faible des diplômés réduit les bénéfices d'être exigeant avec les travailleurs et par conséquent réduit x_T.

Nous constatons que l'impact de la variation du coût de la main-d'œuvre diplômée ou non-diplômée sur les seuils d'exigence respectifs n'est pas symétrique. Une réduction du coût des non-diplômés accroît le bénéfice marginal à traiter un problème additionnel et rend plus rentable l'accroissement du seuil d'exigence des cadres. Au contraire, la réduction du coût salarial des diplômés réduit le bénéfice marginal à être plus exigeant avec les travailleurs et donc le seuil d'exigence de ces derniers.

5 Équilibre

Dans la section précédente, nous avons étudié le comportement d'un principal à options de sortie des employés et distributions des chômeurs données. A présent, nous déterminons les valeurs d'équilibre pour ces variables.

Les options de sortie. La probabilité d'équilibre pour un agent au chômage de retrouver du travail est telle que le flux des entrées au chômage égale le flux des sorties. Sur le marché des non-diplômés, le flux des personnes qui rentrent dans la population des chômeurs est $n_T(1-\alpha+\alpha l_T)$. Il correspond aux travailleurs licenciés $n_T\alpha l_T$ et à ceux qui partent pour des raisons exogènes $n_T(1-\alpha)$. Le flux de sortie[14] est $\lambda_T(N_T-n_T)$. Une proportion λ_T de chômeurs (N_T-n_T) retrouve du travail. Par conséquent, la proportion de chômeurs non-diplômés réembauchés à la fin de la période correspond à :

$$\lambda_T = \frac{n_T(1-\alpha+\alpha l_T)}{(N_T-n_T)} \tag{1.25}$$

La même analyse s'applique au cas des chômeurs diplômés. Ainsi, la probabilité de réembauche garantissant l'équilibre des flux d'entrée et de sortie est :

$$\lambda_C = \frac{n_C(1-\alpha+\alpha l_C)}{(N_C-n_C)} \tag{1.26}$$

Les distributions de productivités. La distribution initiale des non-diplômés, ainsi que celle des diplômés, est divisée en deux distributions stationnaires : celle des agents au sein de l'entreprise et celle des chômeurs. A l'équilibre, les conditions suivantes doivent être vérifiées :

$$U(\theta) = \frac{n_T}{N_T}F(\theta) + \frac{N_T-n_T}{N_T}Q(\theta) \tag{1.27}$$

$$S(\theta) = \frac{n_C}{N_C}\Phi(\theta) + \frac{N_C-n_C}{N_C}P(\theta) \tag{1.28}$$

L'équilibre de cette économie est défini par les équations (1.23) et (1.24) qui déterminent les choix des seuils d'exigence qui maximisent le profit d'un principal à conditions sur le marché du travail données. Les équations (1.25), (1.26), (1.27) et (1.28) garantissent l'équilibre sur le marché du travail.

[14] Nous supposons qu'un agent qui quitte l'organisation ne peut pas retrouver du travail immédiatement, il passe au moins une période au chômage. Il s'agit d'une convention qui n'affecte pas qualitativement nos résultats.

A partir des équations (1.25), (1.26), (1.27) et (1.28), nous constatons qu'à la fois les salaires et la productivité espérée des employés dépendent des conditions sur le marché du travail (l'offre et la demande). L'impact du taux de chômage sur le salaire est bien connu ; un taux de chômage plus faible augmente la probabilité de réembauche. Par conséquent, le licenciement n'est pas une punition forte pour l'employé, ce qui accroît son salaire incitatif.
Le lemme suivant décrit l'impact du taux de chômage sur la productivité espérée des employés :

Lemme 1.7. *l_T est une fonction décroissante du taux de chômage des non-diplômés ($D = T$) et $l_C r_C$ est une fonction décroissante du taux de chômage des diplômés ($D = C$).*

Lorsque la part des employés dans une population d'agents augmente, la distribution de leurs productivités est plus proche de celle qui caractérise la population initiale. Par conséquent, ils sont moins sélectionnés par le principal et la proportion de problèmes transmis à cause de la productivité insuffisante des employés augmente. Ainsi, un taux de chômage plus faible réduit la productivité espérée des employés de la population correspondante.

6 Statique comparative

Dans cette section, nous étudions l'impact de changements sur le marché du travail sur la structure des organisations. Parallèlement, nous discutons les effets des changements organisationnels sur l'emploi (la demande pour les diplômés et les non-diplômés) et les salaires.

L'analyse de cette section est effectuée sous l'Hypothèse 1.7 et les conditions à l'origine de la Proposition 1.1.

Un tableau récapitulatif des résultats de statique comparative est proposé en Annexe 9.14.

6.1 Institutions

Les mesures de régulation du marché du travail affectent les décisions des entreprises. Dans la mesure où nous tenons compte de leur organisation interne, nous pouvons analyser les effets de ce type de mesures sur la population directement concernée par l'intervention, mais aussi sur les autres catégories d'employés ou les propriétaires d'entreprise.

6.1.1 Les indemnités de chômage

L'évolution des indemnités de chômage affecte les options de sortie des employés. Cela modifie le coût salarial des agents de la population affectée par la mesure. Les résultats de la statique comparative qui suit s'appliquent à toute modification institutionnelle qui affecte le coût de la main-d'œuvre. Les conséquences sont cependant différentes si les effets portent à la fois sur le coût et la qualité de la main-d'œuvre, ce qui correspond à une évolution de l'offre de celle-ci par exemple (une discussion plus détaillée de ce point est proposée dans la Section 6.2).

Proposition 1.2. ***Les effets d'une diminution de l'indemnité de chômage des travailleurs (z_T) sur :***

1. *L'organisation :*
   - *Plus de tâches sont déléguées à la fois aux travailleurs ($x_T \nearrow$) et aux cadres ($x_C \nearrow$).*
   - *Le nombre d'employés par cadre augmente.*
2. *L'emploi et les salaires :*
   - *La demande des travailleurs augmente. L'effet sur la demande des cadres est ambigu.*
   - *Le salaire des travailleurs diminue. L'effet sur le salaire des cadres est ambigu.*

Puisque x_C et x_T sont plus élevés, le nombre de travailleurs augmente. Il y a deux effets sur la demande des cadres. Le principal engage d'avantage de travailleurs ce qui devrait augmenter le nombre de cadres nécessaires pour les aider. D'un autre côté, les travailleurs engagés sont plus autonomes, puisque x_T est plus élevé. Par conséquent, l'effet sur l'emploi des cadres est ambigu. Néanmoins, une réduction du coût de la main-d'œuvre non-diplômée peut affecter positivement la main-d'œuvre diplômée à la fois en ce qui concerne leur emploi et leur salaire.

Si nous considérons la possibilité de libre entrée dans cette économie, alors la réduction du coût de la main-d'œuvre non-diplômée n'affecte pas les seuils d'exigence. Cependant, plus d'entreprises entrent sur le marché, ce qui implique une réduction du chômage à la fois pour les non-diplômés et pour les diplômés.

Proposition 1.3. ***Les effets d'une diminution de l'indemnité de chômage des cadres (z_C) sur :***

1. *L'organisation :*
   - *Plus de tâches sont déléguées aux cadres ($x_C \nearrow$), moins de tâches sont déléguées aux travailleurs ($x_T \searrow$).*
   - *Le nombre d'employés par cadre diminue.*
2. *L'emploi et les salaire :*
   - *Plus de cadres sont embauchés, l'effet sur les travailleurs est ambigu.*
   - *Les salaires à la fois des travailleurs et des cadres diminuent.*

Le coût des cadres étant plus faible, le propriétaire en engage un plus grand nombre. L'employeur est plus exigeant avec les cadres (x_C augmente) ; par conséquent la prise de décision formelle (la proportion de problèmes que les cadres sont censés résoudre) est plus décentralisée. Cependant, l'effet sur la productivité réelle est ambigu car $l_C r_C$ augmente fortement. En effet, en plus de l'augmentation due au fait que le principal est à présent plus exigeant avec les cadres, il y a deux effets additionnels allant dans ce même sens. D'abord, la demande pour les cadres augmente significativement[15], ce qui diminue leur taux de chômage et d'après le Lemme 1.7 accroît $l_C r_C$. De plus, $l_C r_C$ est décroissant par rapport au seuil d'exigence des travailleurs. Ici, ce dernier diminue donc $l_C r_C$ augmente. Lorsque ces effets dominent, la proportion de problèmes qui arrivent au niveau du principal est plus importante malgré le choix par l'employeur d'un seuil d'exigence plus élevé pour les cadres. Dès lors, le nombre de tâches que l'organisation peut traiter diminue et le nombre de travailleurs embauchés est réduit en conséquence. Dans ce cas, la diminution du coût des diplômés affecte négativement à la fois l'emploi et les salaires des travailleurs.

6.1.2 Le salaire minimum

Il n'existe pas de résultats univoques quant aux conséquences d'un salaire minimum sur le taux de chômage de la population concernée, ni sur le plan théorique, ni sur celui des études empiriques[16]. Dans l'analyse traditionnelle, si le prix de la main-d'œuvre est augmenté de façon artificielle à un niveau supérieur à celui qui assure l'équilibre du marché, alors il en découle un accroissement du taux de chômage. Néanmoins, il existe des modèles théoriques où l'effet de l'introduction du salaire minimum sur le taux de chômage est bénéfique[17]. Ici, nous analysons les conséquences, sur l'emploi et les salaires des cadres et sur la taille des entreprises, d'une mesure destinée à garantir un salaire minimum aux travailleurs.

Nous introduisons un salaire minimum contraignant pour le propriétaire de l'entreprise (c'est-à-dire $s_T^* < \underline{s}$, où $\underline{s}$ désigne le salaire minimum et s_T^* le salaire

[15] $x_C \nearrow$ et $x_T \searrow$, les deux effets augmentent n_C

[16] Certaines études empiriques montrent l'existence d'un effet nul ou positif (par exemple Card et Krueger (1994) dans le cadre des États Unis, Dickens, Machin et Manning (1999) dans le cadre de la Grande Bretagne), d'autres au contraire, montrent l'existence d'effets négatifs significatifs (par exemple, Kim et Taylor (1995) dans le cadre des États Unis).

[17] Dans les modèles d'appariement par exemple, l'introduction d'un salaire minimum a deux effets : d'un côté puisque le coût de la main-d'œuvre est plus élevé, les entreprises créent moins de postes ; de l'autre côté, le salaire espéré par les demandeurs d'emploi étant plus élevé, ils augmentent leur effort de recherche. Le premier effet diminue la probabilité d'appariement, le second l'augmente. Ainsi, l'effet final sur l'emploi est ambigu. Pour une analyse plus approfondie du sujet, voir Cahuc et Zylberberg (2004).

des travailleurs optimal pour le principal). Dans ce cas, la dérivée du profit $\Pi = n_T(\theta^P - \underline{s} - s_C r_C)$ par rapport à x_T s'écrit :

$$\begin{aligned}\frac{\partial \Pi}{\partial x_T} &= \frac{\partial n_T}{\partial x_T}(\theta^P - \underline{s} - s_C r_C) - n_T(s_C \frac{\partial r_C}{\partial x_T} + r_C \frac{\partial c_C}{\partial x_T}) \\ &= (1 - \frac{\partial l_T}{\partial x_T})\frac{1}{r_P^2}(K + r_P(\frac{c}{\alpha\delta} + (1-\delta)V_U(C)))\end{aligned}$$

où comme dans la Section 4, $K = \displaystyle\int_{\underline{\theta}_C}^{x_C} \frac{\varphi(\theta)(x_C - \theta)\alpha(x_C - \theta - l_C r_C)}{(1-\alpha)r_C + \alpha(x_C - \theta)} d\theta.$

Lemme 1.8. *Soit $\bar{x}_T$ tel que $s_T = \underline{s}$ et $\tilde{x}_T$ tel que $\dfrac{\partial l_T}{\partial x_T} = 1$.*

- *Si $\bar{x}_T < \tilde{x}_T$ alors le seuil d'exigence des travailleurs est $\bar{x}_T$.*
- *Si $\bar{x}_T \geq \tilde{x}_T$ alors le seuil d'exigence des travailleurs est $\tilde{x}_T$.*

L'introduction d'un salaire minimum conduit le principal à devenir plus exigeant avec les travailleurs tant que ceci augmente la proportion de problèmes résolus par ces derniers ($x_T < \tilde{x}_T$). Lorsque l'augmentation de x_T accroît la proportion de problèmes transmis par les travailleurs ($x_T > \tilde{x}_T$), il n'est plus dans l'intérêt du principal d'accroître leur seuil d'exigence.

Proposition 1.4. ***Lorsque $\underline{s}$ augmente ($\underline{s} \geq s_T(\tilde{x}_T)$)***

1. *Le nombre de cadres diminue et leurs salaires sont plus faibles.*
2. *Le nombre de travailleurs par cadre augmente.*
3. *Le nombre de travailleurs diminue.*

L'accroissement du salaire minimum réduit le bénéfice marginal à résoudre un plus grand nombre de problèmes. Dans le cas où $\underline{s} \geq s_T(\tilde{x}_T)$, c'est le seul effet lié à l'introduction du salaire minimum. La conséquence en est une diminution de la demande à la fois des travailleurs et des cadres.

Dans le cas où l'introduction du salaire minimum conduit le principal à augmenter le seuil d'exigence des travailleurs, l'effet sur le chômage des non-diplômés dépend du sens de variation de x_C. En effet, une augmentation de x_T peut entraîner l'accroissement du seuil d'exigence des cadres. Dès lors, l'effet sur l'emploi des travailleurs est ambigu. Néanmoins, dans tous les cas de figure, la demande du principal pour la main-d'œuvre diplômée diminue et le nombre de travailleurs par cadre augmente. Ceci est en adéquation avec les résultats théoriques de Acemoglu et Newman (2002). Dans leur modèle, le travail des cadres consiste à superviser les travailleurs. En effet, les incitations des travailleurs sont assurées par le salaire qui leur est versé et par le degré de supervision qu'ils subissent. Lorsque le nombre d'employés par cadre augmente, la supervision diminue. Des employés qui touchent un salaire supérieur nécessitent moins de supervision et par conséquent moins de cadres seront embauchés dans l'organisation. Dans notre modèle, le salaire minimum affecte à la fois le gain marginal associé à un accroissement du nombre de problèmes résolus par l'organisation et l'allocation des tâches entre les travailleurs et les cadres. Dans leur article, Acemoglu et Newman (2002) fournissent des éléments empiriques en faveur d'un lien positif entre le nombre de travailleurs par cadre et le salaire minimum.

Il est cependant important de souligner que les résultats obtenus ci-dessus s'appliquent aux cas où la hiérarchie de trois niveaux est la forme organisationnelle optimale. Dans le Chapitre 2 de cette thèse, nous montrons que lorsque le principal choisit le nombre de niveaux hiérarchiques, si le coût des travailleurs augmente le propriétaire de l'entreprise peut décider d'adopter une hiérarchie à deux niveaux, où seuls des agents diplômés sont engagés. Dans ce cas, l'introduction du salaire minimum a un effet bénéfique sur les cadres en réduisant leur taux de chômage et en augmentant leurs salaires. Les travailleurs, quant à eux, sont évincés du marché.

6.2 Conditions sur le marché du travail

Nous venons d'analyser les conséquences d'interventions réglementaires sur la structure des organisations. A présent, nous portons notre intérêt sur les implications de changements dans les caractéristiques des employeurs et des évolutions de l'offre de main-d'œuvre.

6.2.1 Propriétaires plus compétents

Nous considérons ici que la productivité des employeurs (θ^P) augmente. Cela peut venir d'une meilleure éducation des propriétaires d'entreprise ou bien d'une amélioration de leur technologie de production.

Proposition 1.5. ***Les effets d'une augmentation de θ^P sur :***

1. *L'organisation :*
   - *Plus de problèmes sont délégués à la fois aux travailleurs ($x_T \nearrow$) et aux cadres ($x_C \nearrow$).*
   - *Le nombre de travailleurs par cadre augmente.*
2. *L'emploi et les salaires :*
   - *D'avantage de travailleurs sont embauchés. L'effet sur le nombre de cadres est ambigu.*
   - *Le salaire des travailleurs augmente. Les effets sur celui des cadres sont ambigus.*

Le principal embauche plus de travailleurs afin d'augmenter le nombre de problèmes traités par l'entreprise. Ceci à son tour augmente la demande pour les agents diplômés nécessaires pour les aider. Néanmoins, dans la mesure où les travailleurs sont plus autonomes l'effet final sur la demande des diplômés est ambigu. Un θ^P plus élevé peut se traduire par une réduction du profit des entreprises. Lorsqu'un principal augmente le nombre de ses employés il exerce une externalité négative sur les autres propriétaires d'entreprise. En effet, lorsque le taux d'emploi augmente cela accroît les options de sortie des agents ($V_U(D) \nearrow$) et augmente ainsi le coût des incitations. Donc, une augmentation de θ^P se traduit par l'adoption d'organisations plus larges mais dont le profit est plus faible. Si nous envisageons la possibilité de libre entrée, le secteur devient plus concentré. Dans ce cas, l'effet final sur le chômage est ambigu. D'un côté, les entreprises embauchent d'avantage, mais de l'autre le nombre d'entreprises sur le marché en question diminue.

6.2.2 Evolutions de l'offre de main-d'œuvre

Les résultats présentés dans cette sous-section sont issus de simulations pour des distributions initiales des productivités ($U(\theta)$ et $S(\theta)$) uniformes.

Les effets d'une augmentation de l'offre de non-diplômés (N_T) sur :

1. L'organisation :
   - Plus de problèmes sont délégués à la fois aux travailleurs ($x_T \nearrow$) et aux cadres ($x_C \nearrow$).
   - Le nombre de travailleurs par cadre augmente.
2. L'emploi et les salaires :
   - La demande de travailleurs augmente et celle de cadres diminue.
   - Le salaire des travailleurs diminue, l'effet sur celui des cadres est ambigu. Le profit augmente.

Quelques exemples de simulations sont proposés en Annexe 9.12

La variation de N_T modifie à la fois le coût et la qualité des travailleurs. Ainsi, l'effet positif sur la proportion de problèmes résolus par les travailleurs est double. D'abord, le principal leur délègue un ensemble de problèmes plus large, c'est-à-dire $x_T \nearrow$. De plus, à partir du Lemme 1.7, nous savons que l'accroissement de l'offre de main-d'œuvre de type $D = T$ réduit l_T. Dans l'entreprise il y a donc plus de travailleurs qui sont plus autonomes (transmettent moins de problèmes à leurs supérieurs). L'effet dominant dans nos simulations est le second et le nombre de cadres engagés diminue.

Dans notre contexte, l'augmentation de l'offre de main-d'œuvre non-diplômée a un effet négatif à la fois sur l'emploi et les salaires des diplômés. Ceci est dû en partie à l'amélioration de la productivité espérée des travailleurs au sein de l'entreprise du fait de la possibilité pour le principal de mieux les sélectionner. Donc, si la distribution des productivités de la main-d'œuvre additionnelle est de moins bonne qualité par exemple, alors la demande de cadres pourrait augmenter.

Les effets d'une augmentation de l'offre de diplômés (N_C) sur :

1. L'organisation :
   - Plus de problèmes sont délégués aux cadres ($x_C \nearrow$) et moins de problèmes sont délégués aux travailleurs ($x_T \searrow$).
   - Les organisations sont plus grandes.
   - Le nombre de travailleurs par cadre diminue.
2. L'emploi et les salaires :
   - La demande de travailleurs et de cadres augmente.
   - Le salaire des cadres diminue, l'effet sur celui des travailleurs est ambigu. Le profit augmente.

Quelques exemples de simulations sont proposés en Annexe 9.13

Comme nous l'avons montré précédemment (dans la Proposition 1.3), la réduction du coût de la main-d'œuvre diplômée peut réduire la demande de travailleurs, du fait de la détérioration de la distribution des cadres. L'augmentation de N_C vient contrecarrer cet effet. En conséquence, la productivité espérée des cadres augmente, ainsi que le nombre de travailleurs engagés.

Ce dernier résultat s'applique tant que la hiérarchie à trois niveaux est l'organisation optimale. En effet, comme nous le montrons dans le Chapitre 2, l'augmentation de N_C, au delà d'un certain seuil, peut conduire à un changement organisationnel et à l'adoption d'organisations constituées de deux niveaux hiérarchiques où seuls des agents diplômés sont engagés. Dans ce cas, les travailleurs sont progressivement évincés du marché et la taille des entreprises, mesurée par le nombre d'employés, diminue.

7 Conclusion

Dans ce chapitre, nous avons proposé des éléments d'analyse permettant de mieux comprendre l'interaction entre les choix organisationnels des entreprises et leur environnement économique. Dans notre modèle, les entreprises organisent leur production en hiérarchies basées sur les compétences et les agents sont affectés aux différents niveaux hiérarchiques en fonction de leur diplôme. A la fois la productivité des travailleurs et des cadres et leurs salaires dépendent des conditions sur le marché du travail. Cela nous a permis de discuter les effets d'évolutions dans ces conditions sur la structure des entreprises en termes de taille, allocation des problèmes, nombre de travailleurs par cadre. De plus, nous avons montré que la prise en compte de l'organisation interne des entreprises permet de mieux comprendre les évolutions sur le marché du travail.

Cependant, notre analyse présente deux limites. Tout d'abord, nous avons considéré que les propriétaires sont homogènes et que toutes les entreprises sont organisées en hiérarchies de trois niveaux. Dans le Chapitre 2, de cette thèse nous analysons le second point. Dans un cadre simplifié, nous endogénéisons le choix du nombre de niveaux hiérarchiques et montrons que l'accroissement de l'offre de diplômés peut conduire à l'adoption d'organisations avec moins de niveaux hiérarchiques et que ce changement organisationnel peut être un des facteurs explicatifs de l'accroissement des inégalités salariales entre diplômés et non-diplômés.

La deuxième restriction, que nous avons imposée, concerne le choix des contrats incitatifs proposés. En effet, nous avons supposé que chaque principal utilise un contrat par poste et que ce contrat est de type salaire d'efficience. Nous pensons qu'élargir l'ensemble des contrats étudiés ainsi qu'inclure la possibilité pour les employeurs de se faire concurrence pour les employés les plus compétents serait une extension intéressante du modèle. Dans la mesure où la forme du contrat optimal peut également dépendre des conditions sur le marché du travail[18], ce serait un élément de discussion supplémentaire pour analyser l'impact de l'environnement de l'entreprise sur sa structure organisationnelle.

[18] Voir par exemple MacLeod et Malcomson (1998) ou le Chapitre 3 de la thèse.

8 Bibliographie

Acemoglu D. et A. Newman (2002), « The Labor Market and Corporate Structure », *European Economic Review*, vol. 46, pp. 1733-1756.

Aghion P. et J. Tirole (1997), « Formal and Real Authority in Organizations », *The Journal of Political Economy*, Vol. 105, pp. 1-29.

Appelbaum E. et R. Batt (1994), *The New American Workplace*, Cornell University Press.

Bolton P. et M. Dewatripont (1994), « The Firm as a Communication Network », *Quarterly Journal of Economics*, vol. 109, pp. 809-839.

Bresnahan T., E. Brynjolfsson et L. Hitt (2002), « Information Technology, Workplace Organization and the Demand for Skilled Labor : Firm-Level Evidence », *Quarterly Journal of Economics*, vol. 117, pp. 339-376.

Cahuc P. et A. Zylberberg (2004), « Labor Economics », *MIT Press*.

Card D. et A.B. Krueger (1994), « Minimum Wages and Employment : A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania », *The American Economic Review*, vol. 84, pp. 772-793.

Caroli E. et J. Van Reenen (2001), « Skill Biased Organizational Change ? Evidence from a panel of British and French Establishments », *Quarterly Journal of Economics*, vol. 116, pp. 1449-1492.

Calvo G. et S. Wellisz (1978), « Supervision, Loss of Control and the Optimal Size of the Firm », *Journal of Political Economy*, vol.86, pp. 943-952.

Dickens R., S. Machin et A. Manning (1999), « The Effects of Minimum Wages on Employment : Theory and Evidence from Britain », *Journal of Labor Economics*, vol. 17, pp. 1-22.

Garicano L. (2000), « Hierarchies and the Organization of Knowledge in Production », *The Journal of Political Economy*, vol. 108, pp. 874-904.

Garicano L. et E. Rossi-Hansberg (2006), « Organization and Inequality in a Knowledge Economy », *Quarterly Journal of Economics*, vol. 121, pp. 1383-1436.

Gibbons R. (1998), « Incentives in Organizations », *The Journal of Economic Perspectives*, Vol. 12, pp. 115-132.

Kim T. et L.J Taylor (1995), « The Employment Effect in Retail Trade of California's 1988 Minimum Wage Increase », *Journal of Business and Economic Statistics*, vol. 13, pp. 175-182.

Kumar K., R. Rajan et L. Zingales (1999), « What Determines firm size ? », NBER working paper 7208.

MacLeod B. et J.M. Malcomson (1998), « Motivation and Markets », *American Economic review*, vol.88, pp. 388-411.

Prendergast C. (1999), « The Provision of Incentives in Firms » *Journal of Economic Literature*, Vol. 37, pp. 7-63.

Radner R. (1992), « Information Processing in Firms and Returns to scale », *Journal of Economic Literature*, vol. 30, pp. 1382-1415.

Radner R. (1993), « The Organization of Decentralized Information processing », *Econometrica*, vol. 61, pp. 1109-1146.

Radner R. et T. Van Zandt (1992), « Information Processing in Firms and Returns to scale », *Annales d'Économie et de Statistique*, vol. 25-26, pp. 265-298.

Rajan R. et J. Wulf (2006), « The Flattening Firm : Evidence from Panel Data on the Changing Nature of Corporate Hierarchies », *The Review of Economics and Statistics*, vol. 88, pp. 759-773.

Roberts J. (2004), « The Modern Firm : Organizational Design for Performance and Growth », *Oxford University Press*.

Shapiro C. et J. Stiglitz (1984), « Equilibrium Unemployment as a Worker Discipline Device », *American Economic Review*, vol. 74, pp. 433-444.

9 Annexe

9.1 Preuve du Lemme 1.3

Le cas des travailleurs. La contrainte d'incitation des travailleurs s'écrit :

$$V(\theta, T) - V_U(\theta, T) \geq \frac{c}{\alpha\delta} \quad (IC)$$

– La contrainte (IC) d'un travailleur $\theta < x_T$.

La différence entre l'utilité inter-temporelle d'un travailleur et d'un chômeur est :

$$V(\theta, T) - V_U(\theta, T) = \frac{s_T - z_T - \theta c}{1 - \alpha\delta(1 - x_T + \theta) + \lambda_T \delta}$$

Elle est obtenue à partir des équations (1.1) et (1.3).

En remplaçant dans la contrainte d'incitation des travailleurs, nous obtenons :

$$V(\theta, T) - V_U(\theta, T) \geq \frac{c}{\alpha\delta} \quad \Leftrightarrow \frac{s_T - z_T - \theta c}{1 - \alpha\delta(1 - x_T + \theta) + \lambda_T \delta} \geq \frac{c}{\alpha\delta}$$

$$\Leftrightarrow s_T \geq \frac{c(1 - \alpha\delta)}{\alpha\delta} + c x_T + z_T + \frac{c\lambda_T}{\alpha}$$

– La contrainte (IC) d'un travailleur $\theta \geq x_T$

La différence entre l'utilité inter-temporelle d'un travailleur et d'un chômeur est :

$$V(\theta, T) - V_U(\theta, T) = \frac{s_T - z_T - x_T c}{1 - \alpha\delta + \lambda_T \delta}$$

Elle est obtenue à partir des équations (1.1) et (1.4).

En remplaçant dans la contrainte d'incitation des travailleurs, nous obtenons :

$$V(\theta,T) - V_U(\theta,T) \geq \frac{c}{\alpha\delta} \quad \Leftrightarrow \frac{s_T - z_T - x_T c}{1-\alpha\delta+\lambda_T\delta} \geq \frac{c}{\alpha\delta}$$

$$\Leftrightarrow s_T \geq \frac{c(1-\alpha\delta)}{\alpha\delta} + cx_T + z_T + \frac{c\lambda_T}{\alpha}$$

Le cas des cadres. Le raisonnement est le même que dans le cas des travailleurs.

– La contrainte (IC) d'un cadre $\theta < x_C$. A partir des équation (1.2) et (1.10) et en tenant compte que dans l'organisation $n = \frac{n_T}{n_C} = \frac{1}{r_C}$ (voir équation (1.17)), nous avons :

$$V(\theta,C) - V_U(\theta,C) \geq \frac{c}{\alpha\delta} \quad \Leftrightarrow \frac{s_C - z_C - Z(\theta)c}{1-\alpha\delta(1-Z(x_C)+Z(\theta))+\lambda_C\delta} \geq \frac{c}{\alpha\delta}$$

$$\Leftrightarrow s_C \geq \frac{c(1-\alpha\delta)}{\alpha\delta} + cZ(x_C) + z_C + \frac{c\lambda_C}{\alpha}$$

– La contrainte (IC) d'un cadre $\theta \geq x_C$. A partir des équation (1.2) et (1.11) et en tenant compte que dans l'organisation $n = \frac{n_T}{n_C} = \frac{1}{r_C}$ (voir équation (1.17)), nous avons :

$$V(\theta,C) - V_U(\theta,C) \geq \frac{c}{\alpha\delta} \quad \Leftrightarrow \frac{s_C - z_C - Z(x_C)c}{1-\alpha\delta+\lambda_C\delta} \geq \frac{c}{\alpha\delta}$$

$$\Leftrightarrow s_C \geq \frac{c(1-\alpha\delta)}{\alpha\delta} + cZ(x_C) + z_C + \frac{c\lambda_C}{\alpha}$$

Si le coût de l'effort ou la probabilité de réembauche dépendent de θ et sous l'hypothèse d'un contrat par niveau hiérarchique, le principal devrait décider non seulement de l'ensemble de problèmes que les agents de chaque niveau hiérarchique devraient résoudre, mais également de l'ensemble d'employés à inciter.

9.2 Les distributions stationnaires

La distribution stationnaire des travailleurs. La distribution stationnaire des travailleurs au début de la période est $F(\theta)$. $\tilde{F}(\theta)$ désigne la fonction de répartition de la distribution modifiée après les départs pour des raisons exogènes $(1-\alpha)$ ou endogènes αl_T. Les travailleurs qui partent sont remplacés par des chômeurs, dont la distribution est $Q(\theta)$.

La condition que vérifient les distributions des travailleurs et des chômeurs non-diplômés à l'état stationnaire est :

$$\underbrace{(1-\alpha+\alpha l_T)}_{\text{Proportion of hired agents}} Q(\theta) + \underbrace{\alpha(1-l_T)}_{\text{Proportion of remaining insiders}} \tilde{F}(\theta) = F(\theta) \tag{1.29}$$

Où $\tilde{F}(u)$ est donné par :

$$\tilde{F}(\theta) = \begin{cases} \dfrac{\displaystyle\int_{\underline{\theta}_T}^{\theta} (1-(x_T-u))f(u)du}{1-l_T} & si \quad \theta \in [\underline{\theta}_T, x_T] \\ \dfrac{F(\theta)-l_T}{1-l_T} & si \quad \theta \in [x_T, \overline{\theta}_T] \end{cases} \tag{1.30}$$

Ainsi, à partir des équations (3.13) et (3.14) nous obtenons :

$$Q(\theta) = \begin{cases} \dfrac{F(\theta)(1-\alpha)+\alpha\displaystyle\int_{\underline{\theta}_T}^{\theta} (x_T-u)f(u)du}{1-\alpha+\alpha l_T} & si \quad u \in [\underline{\theta}_T, x_T] \\ \dfrac{F(\theta)(1-\alpha)+\alpha l_T}{1-\alpha+\alpha l_T} & si \quad \theta \in [x_T, \overline{\theta}_T] \end{cases}$$

La distribution stationnaire des cadres. Nous appliquons la même analyse que précédemment.

- Les agents qui quittent l'organisation :
  - Les cadres licenciés : $\alpha n_C l_C$
  - Les départs exogènes : $(1-\alpha)n_C$
- Les agents qui arrivent dans l'entreprise :
  - Les agents embauchés dans la population des chômeurs diplômés dans le but de remplacer les cadres licenciés ou partis : $(1-\alpha+\alpha l_C)n_C$.

Le lien entre la distribution des cadres ($\Phi(\theta)$) et des chômeurs diplômés ($P(\theta)$) est :

$$P(\theta) = \begin{cases} \dfrac{\Phi(\theta)(1-\alpha) + \alpha \displaystyle\int_{\underline{\theta}_C}^{\theta} \frac{(x_C - u)\varphi(u)}{r_C} du}{1-\alpha+\alpha l_C} & si \quad \theta \in [\underline{\theta}_C; x_C] \\ \dfrac{(1-\alpha)\Phi(\theta) + \alpha l_C}{1-\alpha+\alpha l_C} & si \quad \theta \in [x_C; \overline{\theta}_C] \end{cases}$$

9.3 Preuve du Lemme 1.4

Nous montrons que $Q(\theta) \geq F(\theta)$ pour tout $\theta \in [\underline{\theta}_T, \overline{\theta}_T]$ et $Q(\theta) > F(\theta)$ pour certains θ.

Le cas où $\theta \geq x_T$: A partir de l'équation (1.15), nous savons que :

$$Q(\theta) = \frac{(1-\alpha)F(\theta) + \alpha l_T}{1-\alpha+\alpha l_T} \qquad \text{pour} \qquad \theta \geq x_T$$

Ainsi, nous obtenons :

$$\begin{aligned} Q(\theta) - F(\theta) &= \frac{(1-\alpha)F(\theta) + \alpha l_T}{1-\alpha+\alpha l_T} - F(\theta) \\ &= \frac{\alpha l_T (1 - F(\theta))}{1-\alpha+\alpha l_T} \end{aligned}$$

Il est immédiat de montrer que :

- $\dfrac{\alpha l_T (1 - F(\theta))}{1-\alpha+\alpha l_T} > 0$, pour tout $\theta < \overline{\theta}_T$
- $\dfrac{\alpha l_T (1 - F(\theta))}{1-\alpha+\alpha l_T} = 0$ pour $\theta = \overline{\theta}_T$

Le cas où $\theta < x_T$: A partir de l'équation (1.15), nous savons que :

$$Q(\theta) = \frac{(1-\alpha)F(\theta) + \alpha \displaystyle\int_{\underline{\theta}_T}^{\theta} (x_T - u) f(u) du}{1-\alpha+\alpha l_T} \qquad \text{pour} \qquad \theta < x_T$$

Par conséquent :

$$\begin{aligned} Q(\theta) - F(\theta) &= \frac{(1-\alpha)F(\theta) + \alpha \displaystyle\int_{\underline{\theta}_T}^{\theta} (x_T - u) f(u) du}{1 - \alpha + \alpha l_T} - F(\theta) \\ &= \frac{\alpha \displaystyle\int_{\underline{\theta}_T}^{\theta} (x_T - u - l_T) f(u) du}{1 - \alpha + \alpha l_T} \end{aligned}$$

$\displaystyle\int_{\underline{\theta}_T}^{\theta} (x_T - u - l_T) f(u) du$ augmente avec θ lorsque $\theta < x_T - l_T$ et diminue avec θ lorsque $\theta > x_T - l_T$. Donc, si $\displaystyle\int_{\underline{\theta}_T}^{\theta} (x_T - u - l_T) f(u) du \geq 0$ pour $\theta = \underline{\theta}_T$ et $\theta = x_T$, alors il sera également positif pour tout $\theta \in]\underline{\theta}_T, x_T[$.

- Pour $\theta = x_T$, $\displaystyle\int_{\underline{\theta}_T}^{x_T} (x_T - u - l_T) f(u) du = (1 - F(x_T)) l_T > 0$.
- Pour $\theta = \underline{\theta}_T$, $\displaystyle\int_{\underline{\theta}_T}^{\underline{\theta}_T} (x_T - u - l_T) f(u) du = 0$.

Ainsi, nous avons montré que $Q(\theta) > F(\theta)$ pour $\theta \in]\underline{\theta}_T, \overline{\theta}_T[$ et $Q(\theta) = F(\theta)$ pour $\theta = \underline{\theta}_T$ et $\theta = \overline{\theta}_T$.

La même analyse s'applique à la distribution des diplômés.

9.4 La dérivée de $(l_C r_C)$ par rapport à x_T

Par définition $l_C r_C = \bar{\Phi}(x_C)$. Par conséquent, (à partir du Lemme 1.1) :

$$l_C r_C = \int_{\underline{\theta}_C}^{x_C} (x_C - \theta) \varphi(\theta) d\theta$$

A partir de l'équation (1.16) nous obtenons :

$$\varphi(\theta) = \frac{(1-\alpha) r_C + \alpha r_C l_C}{(1-\alpha) r_C + \alpha (x_C - \theta)} p(\theta) \qquad \text{pour} \qquad \theta < x_C$$

Par conséquent :

$$l_C r_C = \int_{\underline{\theta}_C}^{x_C} \frac{p(\theta)(x_C - \theta)((1-\alpha) r_C + \alpha (l_C r_C))}{(1-\alpha) r_C + \alpha (x_C - \theta)} d\theta$$

A présent, nous pouvons en écrire la dérivée par rapport à x_T, à distribution des chômeurs diplômés ($p(\theta)$) donnée.

$$\begin{aligned}
\frac{\partial l_C r_C}{\partial x_T} &= \frac{\partial\Big(\int_{\underline{\theta}_C}^{x_C} \frac{p(\theta)(x_C-\theta)((1-\alpha)r_C+\alpha(l_C r_C))}{(1-\alpha)r_C+\alpha(x_C-\theta)} d\theta\Big)}{\partial x_T} \\
&= \frac{\partial l_C r_C}{\partial x_T}\int_{\underline{\theta}_C}^{x_C} \frac{p(\theta)\alpha(x_C-\theta)}{(1-\alpha)r_C+\alpha(x_C-\theta)} d\theta + (1-\alpha)\frac{\partial r_C}{\partial x_T}\int_{\underline{\theta}_C}^{x_C} \frac{p(\theta)(x_C-\theta)\alpha(x_C-\theta-l_C r_C)}{((1-\alpha)r_C+\alpha(x_C-\theta))^2} d\theta \\
&= \frac{\partial l_C r_C}{\partial x_T}\frac{\alpha l_C r_C}{(1-\alpha)r_C+\alpha l_C r_C} + (1-\alpha)\frac{\partial r_C}{\partial x_T}\int_{\underline{\theta}_C}^{x_C} \frac{p(\theta)(x_C-\theta)\alpha(x_C-\theta-l_C r_C)}{((1-\alpha)r_C+\alpha(x_C-\theta))^2} d\theta \\
&= -\frac{(1-\frac{\partial l_T}{\partial x_T})}{r_C}\int_{\underline{\theta}_C}^{x_C} \frac{p(\theta)(x_C-\theta)\alpha(x_C-\theta-(l_C r_C))((1-\alpha)r_C+\alpha(l_C r_C))}{((1-\alpha)r_C+\alpha(x_C-\theta))^2} d\theta \\
&= -\frac{(1-\frac{\partial l_T}{\partial x_T})}{r_C}\int_{\underline{\theta}_C}^{x_C} \frac{\varphi(\theta)(x_C-\theta)\alpha(x_C-\theta-(l_C r_C))}{((1-\alpha)r_C+\alpha(x_C-\theta))} d\theta \\
&= -\frac{(1-\frac{\partial l_T}{\partial x_T})}{r_C} K
\end{aligned}$$

où $K = \displaystyle\int_{\underline{\theta}_C}^{x_C} \frac{\varphi(\theta)(x_C-\theta)\alpha(x_C-\theta-(l_C r_C))}{((1-\alpha)r_C+\alpha(x_C-\theta))} d\theta.$

A présent, nous allons montrer que $K > 0$.

$\dfrac{(x_C-\theta)}{((1-\alpha)r_C+\alpha(x_C-\theta))}$ est décroissant par rapport à θ, donc dans l'intégrale K les termes positifs sont plus fortement pondérés. Par conséquent, si l'expression de l'intégrale est positive dans le cas où les termes $\varphi(\theta)(x_C-\theta-(l_C r_C))$ ont les mêmes poids, elle le sera également lorsque les termes positifs sont plus fortement pondérés. Si tous les termes ont les mêmes poids nous avons :

$$\int_{\underline{\theta}_C}^{x_C} \varphi(\theta)\alpha(x_C-\theta-(l_C r_C))d\theta = \alpha(l_C r_C)(1-\Phi(x_C)) > 0$$

Donc dans le mesure où dans K les termes positifs sont plus fortement pondérés $K > 0$.

9.5 Preuve du Lemme 1.5

La preuve est détaillée pour le cas de l_T. Le même raisonnement s'applique à $(l_C r_C)$.

Par définition :

$$l_T = \int_{\underline{\theta}_T}^{x_T} (x_T - \theta) f(\theta) d\theta$$

A partir de l'équation (1.15) nous obtenons :

$$f(\theta) = \frac{q(\theta)(1 - \alpha + \alpha l_T)}{1 - \alpha + \alpha(x_T - \theta)} \qquad \text{pour} \qquad \theta < x_T$$

Par conséquent :

$$l_T = \int_{\underline{\theta}_T}^{x_T} \frac{(x_T - \theta) q(\theta)(1 - \alpha + \alpha l_T)}{(1 - \alpha + \alpha(x_T - \theta))} d\theta$$

Dans un premier temps, nous dérivons l_T par rapport à x_T à distribution des chômeurs non-diplômés ($q(theta)$) donnée. Ensuite, nous montrons que $\frac{\partial l_T}{\partial x_T} > 1$ pour certaines valeurs de x_T.

Dans la mesure où le principal maximise son profit à distribution des chômeurs non-diplômés ($q(\theta)$) donnée, nous avons :

$$\begin{aligned}
\frac{\partial l_T}{\partial x_T} &= \frac{\partial \left(\int_{\underline{\theta}_T}^{x_T} \frac{q(\theta)(x_T - \theta)(1 - \alpha + \alpha l_T)}{1 - \alpha + \alpha(x_T - \theta)} d\theta \right)}{\partial x_T} \\
&= \int_{\underline{\theta}_T}^{x_T} \frac{q(\theta)(1 - \alpha + \alpha l_T)(1 - \alpha)}{(1 - \alpha + \alpha(x_T - \theta))^2} d\theta + \alpha \frac{\partial l_T}{\partial x_T} \int_{\underline{\theta}_T}^{x_T} \frac{q(\theta)(x_T - \theta)}{1 - \alpha + \alpha(x_T - \theta)} d\theta \\
&= \int_{\underline{\theta}_T}^{x_T} \frac{q(\theta)(1 - \alpha + \alpha l_T)(1 - \alpha)}{(1 - \alpha + \alpha(x_T - \theta))^2} d\theta + \frac{\partial l_T}{\partial x_T} \frac{\alpha l_T}{1 - \alpha + \alpha l_T} \\
&= \int_{\underline{\theta}_T}^{x_T} \frac{q(\theta)(1 - \alpha + \alpha l_T)^2}{(1 - \alpha + \alpha(x_T - \theta))^2} d\theta \\
&= \int_{\underline{\theta}_T}^{x_T} \frac{f(\theta)(1 - \alpha + \alpha l_T)}{(1 - \alpha + \alpha(x_T - \theta))} d\theta
\end{aligned}$$

Afin de montrer qu'il existe des valeurs de x_T pour lesquelles :

$$\int_{\underline{\theta}_T}^{x_T} \frac{f(\theta)(1 - \alpha + \alpha l_T)}{1 - \alpha + \alpha(x_T - \theta)} d\theta > 1$$

nous considérons le cas où $x_T = \overline{\theta}_T$.

$$\begin{aligned}\int_{\underline{\theta}_T}^{\overline{\theta}_T} \frac{f(\theta)(1-\alpha+\alpha l_T)}{1-\alpha+\alpha(\overline{\theta}_T-\theta)}d\theta > 1 \quad &\Leftrightarrow \int_{\underline{\theta}_T}^{\overline{\theta}_T} \frac{f(\theta)(1-\alpha+\alpha l_T)}{1-\alpha+\alpha(\overline{\theta}_T-\theta)}d\theta - \int_{\underline{\theta}_T}^{\overline{\theta}_T} f(\theta)d\theta > 0 \\ &\Leftrightarrow \int_{\underline{\theta}_T}^{\overline{\theta}_T} \frac{f(\theta)(\overline{\theta}_T-\theta-l_T)}{1-\alpha+\alpha(\overline{\theta}_T-\theta)}d\theta < 0\end{aligned}$$

$\dfrac{1}{1-\alpha+\alpha(\overline{\theta}_T-\theta)}$ est croissant en θ, donc les termes négatifs sont plus fortement pondérés.

Si tous les termes ont les mêmes poids alors : $\int_{\underline{\theta}_T}^{\overline{\theta}_T} f(\theta)(\overline{\theta}_T-\theta-l_T)d\theta = l_T - l_T = 0$. Puisque les termes négatifs sont plus fortement pondérés :

$$\int_{\underline{\theta}_T}^{\overline{\theta}_T} \frac{f(\theta)(\overline{\theta}_T-\theta-l_T)}{1-\alpha+\alpha(\overline{\theta}_T-\theta)}d\theta < 0$$

En utilisant le même raisonnement nous obtenons :

$$\frac{\partial(l_C r_C)}{\partial x_C} = \int_{\underline{\theta}_C}^{x_C} \frac{\varphi(\theta)((1-\alpha)r_C+\alpha(l_C r_C))}{((1-\alpha)r_C+\alpha(x_C-\theta))}d\theta$$

et pouvons montrer que :

$$\int_{\underline{\theta}_C}^{x_C} \frac{\varphi(\theta)((1-\alpha)r_C+\alpha(l_C r_C))}{((1-\alpha)r_C+\alpha(x_C-\theta))}d\theta > 1 \text{ si } x_C = \overline{\theta}_C$$

9.6 Discussion de l'Hypothèse 1.7

Dans l'Annexe 9.5, nous avons montré que :

$$\frac{\partial l_T}{\partial x_T} = \int_{\underline{\theta}_T}^{x_T} \frac{q(\theta)(1-\alpha+\alpha l_T)^2}{(1-\alpha+\alpha(x_T-\theta))^2}d\theta$$

En dérivant cette expression par rapport à x_T à distribution des chômeurs non-diplômés ($q(\theta)$) donnée et après simplification nous obtenons :

$$\frac{\partial^2 l_T}{\partial x_T^2} = \frac{f(x_T)(1-\alpha+\alpha l_T)}{1-\alpha} + \frac{2\alpha}{1-\alpha+\alpha l_T}\Big(\frac{\partial l_T}{\partial x_T}\Big)^2 - 2\alpha \int_{\underline{\theta}_T}^{x_T} \frac{f(\theta)(1-\alpha+\alpha l_T)}{(1-\alpha+\alpha(x_T-\theta))^2} d\theta$$

Ce qui nous intéresse c'est le signe de cette dérivée, nous pouvons donc écrire :

$$\frac{\partial^2 l_T}{\partial x_T^2} \propto \Big[f(x_T) + 2\alpha(1-\alpha)\Big(\Big(\int_{\underline{\theta}_T}^{x_T} \frac{f(\theta)}{(1-\alpha+\alpha(x_T-\theta))} d\theta\Big)^2 - \int_{\underline{\theta}_T}^{x_T} \frac{f(\theta)}{(1-\alpha+\alpha(x_T-\theta))^2} d\theta\Big)\Big] \tag{1.31}$$

Nous souhaitons montrer qu'il existe des distributions pour lesquelles $\frac{\partial^2 l_T}{\partial x_T^2} > 0$ pour tout x_T.

Le côté droit de (1.31) est minoré par :

$$f(x_T) - 2\alpha(1-\alpha) \int_{\underline{\theta}_T}^{x_T} \frac{f(\theta)}{(1-\alpha+\alpha(x_T-\theta))^2} d\theta$$

Si $f(\theta)$ est croissant en θ, vérifié pour une grande majorité de distributions, $\frac{\partial u(\theta)}{\partial \theta} \geq 0$, alors :

$$f(x_T) - 2\alpha(1-\alpha) \int_{\underline{\theta}_T}^{x_T} \frac{f(\theta)}{(1-\alpha+\alpha(x_T-\theta))^2} d\theta \geq f(x_T)(1 - 2\alpha \frac{x_T - \underline{\theta}_T}{1-\alpha+\alpha(x_T-\underline{\theta}_T)})$$

Cette dernière expression est positive pour tout x_T si $\overline{\theta}_T - \underline{\theta}_T \leq \frac{1-\alpha}{\alpha}$.

Donc $\frac{\partial^2 l_T}{\partial x_T^2} > 0$ pour tout x_T, si la distribution initiale vérifie :

$$\frac{\partial u(\theta)}{\partial \theta} \geq 0 \qquad et \qquad \overline{\theta}_T - \underline{\theta}_T \leq \frac{1-\alpha}{\alpha}$$

Cependant, il s'agit ici de conditions suffisantes fortes et l'ensemble des distributions initiales pour lesquelles la dérivée seconde de l_T est positive est plus large.

9.7 Preuve de la Proposition 1.1

Nous procédons en trois étapes :

- Dans un premier temps, nous calculons les dérivées secondes du profit par rapport aux seuils d'exigence et étudions la nature de l'interaction entre les seuils d'exigence (substituts ou compléments).
- Ensuite, nous donnons les conditions sous lesquelles la complémentarité ou la substituabilité des seuils d'exigence n'est pas trop forte. Ces conditions garantissent que dans la statique comparative les effets indirects sont systématiquement dominés. Nous montrons que, sous ces conditions, les conditions de second ordre du programme du principal sont vérifiées.
- Enfin, nous démontrons les résultats de la Proposition 1.1.

9.7.1 Compléments ou substituts

Nous calculons donc, pour les valeurs d'équilibre, les dérivées secondes du profit par rapport aux seuils d'exigence.

$$\frac{\partial^2\Pi}{\partial x_T^2} = -\frac{\partial^2(l_C r_C)}{\partial x_T^2}\Pi - \frac{\partial^2 l_T}{\partial x_T^2}(\frac{c}{\alpha\delta} + (1-\delta)V_U(C))$$
$$\frac{\partial^2\Pi}{\partial x_C^2} = -\frac{\partial^2(l_C r_C)}{\partial x_C^2}\Pi$$
$$\frac{\partial^2\Pi}{\partial x_T\partial x_C} = -\frac{\partial^2(l_C r_C)}{\partial x_T\partial x_C}\Pi$$

Ainsi, nous avons :

$$\frac{\partial x_T}{\partial x_C} = -\frac{\frac{\partial^2\Pi}{\partial x_T\partial x_C}}{\frac{\partial^2\Pi}{\partial x_T^2}} = -\frac{\frac{\partial^2(l_C r_C)}{\partial x_T\partial x_C}}{\frac{\partial^2(l_C r_C)}{\partial x_C^2}}$$
$$\frac{\partial x_C}{\partial x_T} = -\frac{\frac{\partial^2\Pi}{\partial x_T\partial x_C}}{\frac{\partial^2\Pi}{\partial x_C^2}} = -\frac{\frac{\partial^2(l_C r_C)}{\partial x_T\partial x_C}}{\frac{\partial^2(l_C r_C)}{\partial x_T^2} + \frac{\partial^2 l_T}{\partial x_T^2}(\frac{c}{\alpha\delta} + (1-\delta)V_U(C))\frac{1}{\Pi}}$$

Comme ceci a été montré précédemment :

$$\frac{\partial(l_C r_C)}{\partial x_C} = \int_{\underline{\theta}_C}^{x_C} \frac{p(\theta)((1-\alpha)r_C + \alpha(l_C r_C))^2}{((1-\alpha)r_C + \alpha(x_C - \theta))^2}$$

Donc après simplification nous obtenons :

$$\begin{aligned}\frac{\partial^2(l_C r_C)}{\partial x_C \partial x_T} = \;& 2\alpha \frac{\partial(l_C r_C)}{\partial x_T} \int_{\underline{\theta}_C}^{x_C} \frac{\varphi(\theta)}{(1-\alpha)r_C + \alpha(x_C - \theta)} d\theta \\ & -2\alpha(1-\alpha)(1 - \frac{\partial l_T}{\partial x_T}) \int_{\underline{\theta}_C}^{x_C} \frac{\varphi(\theta)(x_C - \theta - (l_C r_C))}{((1-\alpha)r_C + \alpha(x_C - \theta))^2} d\theta\end{aligned}$$

Cette dernière expression peut être positive ou négative en fonction des valeurs des paramètres. Si $\frac{\partial^2(l_C r_C)}{\partial x_C \partial x_T} < 0$, alors les seuils de performance sont des compléments, c'est-à-dire $\frac{\partial^2 \Pi}{\partial x_T \partial x_C} > 0$. Si $\frac{\partial^2(l_C r_C)}{\partial x_C \partial x_T} > 0$, alors les seuils d'exigence sont des substituts, c'est-à-dire $\frac{\partial^2 \Pi}{\partial x_T \partial x_C} < 0$.

9.7.2 Effets indirects dominés et conditions de second ordre

Faible complémentarité. Il existe deux conditions pour sous lesquelles les effets de complémentarité ne sont pas trop fort à l'équilibre. Ces conditions s'écrivent :

$$\frac{\partial x_C}{\partial x_T} < \frac{(1 - \frac{\partial l_T}{\partial x_T})r_P + \frac{\partial(l_C r_C)}{\partial x_T} r_C}{(1 - \frac{\partial(l_C r_C)}{\partial x_C})r_C} \tag{1.32}$$

Cela garantit que le nombre de cadres est décroissant par rapport à x_T.

$$\frac{\partial x_T}{\partial x_C} < \frac{(1 - \frac{\partial(l_C r_C)}{\partial x_C})r_C}{(1 - \frac{\partial l_T}{\partial l_T})r_P + \frac{\partial(l_C r_C)}{\partial x_T} r_C} \tag{1.33}$$

Cela garantit que le nombre de cadres est croissant par rapport à x_C.

Faible substituabilité Dans le cas où $\Pi_{x_T,x_C} < 0$, les deux conditions suivantes garantissent que les effets directs dominent dans notre analyse de statique comparative.

$$-\frac{\partial x_C}{\partial x_T} < -\frac{\frac{\partial(l_C r_C)}{\partial x_T}}{(1-\frac{\partial(l_C r_C)}{\partial x_C})} \tag{1.34}$$

Cela garantit que le nombre de travailleurs est croissant par rapport à x_T.

$$-\frac{\partial x_T}{\partial x_C} < \frac{(1-\frac{\partial(l_C r_C)}{\partial x_C})}{-\frac{\partial(l_C r_C)}{\partial x_T}} \tag{1.35}$$

Cela garantit que le nombre de travailleurs est croissant par rapport à x_C.

9.7.2.1 Conditions de second ordre

$$J = \begin{pmatrix} \Pi_{x_T,x_T} & \Pi_{x_T,x_C} \\ \Pi_{x_C,x_T} & \Pi_{x_C,x_C} \end{pmatrix}$$

Les conditions de second ordre sont vérifiées si :

$$\begin{cases} \Pi_{x_T,x_T} + \Pi_{x_C,x_C} < 0 \\ |J| > 0 \end{cases}$$

- $\Pi_{x_T,x_T} + \Pi_{x_C,x_C} < 0$

$$\Pi_{x_T,x_T} + \Pi_{x_C,x_C} = -\frac{\partial^2(l_C r_C)}{\partial x_T^2}\Pi - \frac{\partial^2 l_T}{\partial x_T^2}(\frac{c}{\alpha\delta} + (1-\delta)V_U(C)) - \frac{\partial^2(l_C r_C)}{\partial x_C^2}\Pi < 0$$

Nous étudions des situations où $\frac{\partial^2 l_T}{\partial x_T^2} > 0$ et $\frac{\partial^2(l_C r_C)}{\partial x_C^2} > 0$.

- $|J| > 0$

$$\begin{aligned} |J| &= \Pi_{x_T,x_T}\Pi_{x_C,x_C} - (\Pi_{x_C,x_T})^2 \\ &= \Pi_{x_T,x_T}\Pi_{x_C,x_C}(1 - \frac{\partial x_T}{\partial x_C}\frac{\partial x_C}{\partial x_T}) \\ &> 0 \end{aligned}$$

Nous pouvons montrer que, sous les conditions (1.32) et (1.33), $|J| > 0$. Lorsque les conditions sont vérifiées avec égalité et après simplification nous avons :

$$\begin{aligned} |J| &= \Pi_{x_T,x_T}\Pi_{x_C,x_C}\left(1 - \frac{(1-\frac{\partial(l_C r_C)}{\partial x_C})r_C}{(1-\frac{\partial l_T}{\partial x_T})r_P + \frac{\partial(l_C r_C)}{\partial x_T}r_C}\frac{(1-\frac{\partial l_T}{\partial x_T})r_P + \frac{\partial(l_C r_C)}{\partial x_T}r_C}{(1-\frac{\partial(l_C r_C)}{\partial x_C})r_C}\right) \\ &= 0 \end{aligned}$$

Donc, sous les conditions de faible complémentarité $|J| > 0$.

Le même raisonnement peut s'appliquer pour les conditions (1.34) et (1.35).

9.7.3 Preuve de la Proposition 1.1

Les conditions de premier ordre définissent le système suivant :

$$\begin{cases} \Pi_{x_T}(x_T, x_C; \theta^P, c...) = 0 \\ \Pi_{x_C}(x_T, x_C; \theta^P, c...) = 0 \end{cases}$$

– Variation de x_T de performance par rapport à θ^P

Pour déterminer le sens de variation de x_i, en fonction du paramètre θ^P, nous commençons par écrire :

$$\begin{pmatrix} \Pi_{x_T,x_T} & \Pi_{x_T,x_C} \\ \Pi_{x_C,x_T} & \Pi_{x_C,x_C} \end{pmatrix} \begin{pmatrix} \frac{\partial x_T}{\theta^P} \\ \frac{\partial x_C}{\theta^P} \end{pmatrix} = - \begin{pmatrix} \Pi_{x_T,\theta^P} \\ \Pi_{x_C,\theta^P} \end{pmatrix}$$

En utilisant la règle de Cramer, nous avons :

$$\frac{\partial x_i}{\partial \theta^P} = -\frac{|J_i|}{|J|}$$

où J_i est la matrice obtenue en remplaçant la i^{me} colonne du Jacobian (J) par le vecteur $(\Pi_{x_T,\theta^P}, \Pi_{x_C,\theta^P})^T$. Donc dans le cas de x_T ce la donne :

$$\frac{\partial x_T}{\partial \theta^P} = -\frac{\begin{vmatrix} \Pi_{x_T,\theta^P} & \Pi_{x_T,x_C} \\ \Pi_{x_C,\theta^P} & \Pi_{x_C,x_C} \end{vmatrix}}{\begin{vmatrix} \Pi_{x_T,x_T} & \Pi_{x_T,x_C} \\ \Pi_{x_C,x_T} & \Pi_{x_C,x_C} \end{vmatrix}}$$

$$\begin{aligned} \frac{\partial x_T}{\partial \theta^P} &= -\frac{\Pi_{x_T,\theta^P}\Pi_{x_C,x_C} - \Pi_{x_T,x_C}\Pi_{x_C,\theta^P}}{|J|} \\ &= -\frac{\Pi_{x_C,x_C}}{|J|}\Big(\Pi_{x_T,\theta^P} + \frac{\partial x_C}{\partial x_T}\Pi_{x_C,\theta^P}\Big) \end{aligned}$$

Comme ceci a déjà été montré $|J| > 0$ et $\Pi_{x_C,x_C} < 0$. $\Pi_{x_C,\theta^P} = \Big(1 - \frac{\partial(l_C r_C)}{\partial x_C}\Big)\frac{1}{r_P} \Rightarrow \Pi_{x_C,\theta^P}\ |_{x_C=x_m^*} > 0$, $\Pi_{x_T,\theta^P} = -\frac{\partial(l_C r_C)}{\partial x_T}\frac{1}{r_P} \Rightarrow \Pi_{x_T,\theta^P}\ |_{x_T=x_T^*} > 0$.

$$\frac{\partial x_T}{\partial \theta^P} = -\frac{\Pi_{x_C,x_C}}{|J|}\frac{1}{r_P}\Big[-\frac{\partial(l_C r_C)}{\partial x_T} + \frac{\partial x_C}{\partial x_T}\Big(1 - \frac{\partial(l_C r_C)}{\partial x_C}\Big)\Big]$$

Si $\frac{\partial x_C}{\partial x_T} > 0$, il est facile de voir que $\frac{\partial x_T}{\partial \theta^P} > 0$.

Si $\frac{\partial x_C}{\partial x_T} < 0$, alors $\frac{\partial x_T}{\partial \theta^P} > 0$ sous la condition (1.34).

– Variation de x_C de performance par rapport à θ^P

$$\begin{aligned} \frac{\partial x_C}{\partial \theta^P} &= -\frac{\Pi_{x_T,x_T}\Pi_{x_C,\theta^P} - \Pi_{x_T,x_C}\Pi_{x_T,\theta^P}}{|J|} \\ &= -\frac{\Pi_{x_T,x_T}}{|J|}(\Pi_{x_C,\theta^P} + \frac{\partial x_T}{\partial x_C}\Pi_{x_T,\theta^P}) \\ &= -\frac{\Pi_{x_T,x_T}}{|J|}\frac{1}{r_P}\Big[\Big(1 - \frac{\partial(l_C r_C)}{\partial x_C}\Big) - \frac{\partial(l_C r_C)}{\partial x_T}\frac{\partial x_T}{\partial x_C}\Big] \end{aligned}$$

$\Pi_{x_T,x_T} < 0$.

Si $\frac{\partial x_C}{\partial x_T} > 0$ il est facile de voir que $\frac{\partial x_C}{\partial \theta^P} > 0$.

Si $\frac{\partial x_C}{\partial x_T} < 0$, alors $\frac{\partial x_T}{\partial \theta^P} > 0$ sous la condition (1.35).

– Variation de x_T par rapport à $V_U(T)$

$$\begin{aligned}\frac{\partial x_T}{\partial V_U(T)} &= -\frac{\Pi_{x_T,V_U(T)}\Pi_{x_C,x_C} - \Pi_{x_T,x_C}\Pi_{x_C,V_U(T)}}{|J|} \\ &= -\frac{\Pi_{x_C,x_C}}{|J|}(\Pi_{x_T,V_U(T)} + \frac{\partial x_C}{\partial x_T}\Pi_{x_C,V_U(T)})\end{aligned}$$

où $\Pi_{x_C,V_U(T)} = -\Big(1 - \dfrac{\partial(l_C r_C)}{\partial x_C}\Big)\dfrac{(1-\delta)}{r_P} < 0$ et $\Pi_{x_T,V_U(T)} = \dfrac{\partial(l_C r_C)}{\partial x_T}\dfrac{(1-\delta)}{r_P} < 0.$

Si $\dfrac{\partial x_C}{\partial x_T} > 0$, il est facile de voir que $\dfrac{\partial x_T}{\partial V_U(T)} < 0$.

Si $\dfrac{\partial x_C}{\partial x_T} < 0$, alors $\dfrac{\partial x_T}{\partial V_U(T)} < 0$ sous la condition (1.34).

– Variation des x_C par rapport à $V_U(T)$

$$\begin{aligned}\frac{\partial x_C}{\partial V_U(T)} &= -\frac{\Pi_{x_T,x_T}\Pi_{x_C,V_U(T)} - \Pi_{x_T,x_C}\Pi_{x_T,V_U(T)}}{|J|} \\ &= -\frac{\Pi_{x_T,x_T}}{|J|}(\Pi_{x_C,V_U(T)} + \frac{\partial x_T}{\partial x_C}\Pi_{x_T,V_U(T)}) \\ &= \frac{\Pi_{x_T,x_T}}{|J|}\frac{(1-\delta)}{r_P}\Big[\Big(1 - \frac{\partial(l_C r_C)}{\partial x_C}\Big) - \frac{\partial(l_C r_C)}{\partial x_T}\frac{\partial x_T}{\partial x_C}\Big]\end{aligned}$$

Si $\dfrac{\partial x_C}{\partial x_T} > 0$, il est facile de voir que $\dfrac{\partial x_C}{\partial V_U(T)} < 0$.

Si $\dfrac{\partial x_C}{\partial x_T} < 0$, alors $\dfrac{\partial x_T}{\partial V_U(T)} < 0$ sous la condition (1.35).

– Variation de x_T avec $V_U(C)$

$$\begin{aligned}\Pi_{x_T,V_U(C)} &= \frac{\partial(l_C r_C)}{\partial x_T}\frac{r_C(1-\delta)}{r_P} + (1-\delta)(1 - \frac{\partial l_T}{\partial x_T}) \\ &= \frac{(1-\delta)(1-\frac{\partial l_T}{\partial x_T})}{r_P}\Big(r_P - \int_{\underline{\theta}_C}^{x_C}\frac{\alpha(x_C-\theta)(x_C-\theta-(l_C r_C))\varphi(\theta)}{(1-\alpha)r_C + \alpha(x_C-\theta)}d\theta\Big) \\ &= \frac{(1-\delta)(1-\frac{\partial l_T}{\partial x_T})}{r_P}\Big(1 - x_C + l_C r_C - \int_{\underline{\theta}_C}^{x_C}\frac{\alpha(x_C-\theta)(x_C-\theta-(l_C r_C))\varphi(\theta)}{(1-\alpha)r_C + \alpha(x_C-\theta)}d\theta\Big) \\ &= \frac{(1-\delta)(1-\frac{\partial l_T}{\partial x_T})}{r_P}\Big(1 - x_C + \int_{\underline{\theta}_C}^{x_C}\frac{(x_C-\theta)((1-\alpha)r_C + \alpha(l_C r_C))\varphi(\theta)}{(1-\alpha)r_C + \alpha(x_C-\theta)}d\theta\Big) \\ &> 0\end{aligned}$$

$$\begin{aligned}\Pi_{x_C,V_U(C)} &= -\Big(1 - \frac{\partial(l_C r_C)}{\partial x_C}\Big)\frac{(1-\delta)r_C}{r_P} \\ &< 0\end{aligned}$$

$$\begin{aligned}
\frac{\partial x_T}{\partial V_U(C)} &= -\frac{\Pi_{x_T,V_U(C)}\Pi_{x_C,x_C} - \Pi_{x_T,x_C}\Pi_{x_C,V_U(C)}}{|J|} \\
&= -\frac{\Pi_{x_C,x_C}}{|J|}\Big(\Pi_{x_T,V_U(C)} + \frac{\partial x_C}{\partial x_T}\Pi_{x_C,V_U(C)}\Big) \\
&= -\frac{\Pi_{x_C,x_C}}{|J|}\Big[\frac{\partial(l_C r_C)}{\partial x_T}\frac{r_C(1-\delta)}{r_P} + (1-\delta)\Big(1-\frac{\partial l_T}{\partial x_T}\Big) - \frac{\partial x_C}{\partial x_T}\Big(1-\frac{\partial(l_C r_C)}{\partial x_C}\Big)\frac{(1-\delta)r_C}{r_P}\Big]
\end{aligned}$$

Si $\frac{\partial x_C}{\partial x_T} > 0$, $\frac{\partial x_T}{\partial V_U(C)} > 0$ sous la condition (1.32).

Si $\frac{\partial x_C}{\partial x_T} < 0$ il est facile de voir que $\frac{\partial x_T}{\partial V_U(C)} > 0$.

– Variation des x_C avec $V_U(C)$

$$\begin{aligned}
\frac{\partial x_C}{\partial V_U(C)} &= -\frac{\Pi_{x_T,x_T}\Pi_{x_C,V_U(C)} - \Pi_{x_T,x_C}\Pi_{x_T,V_U(C)}}{|J|} \\
&= -\frac{\Pi_{x_T,x_T}}{|J|}\Big(\Pi_{x_C,V_U(C)} + \frac{\partial x_T}{\partial x_C}\Pi_{x_T,V_U(C)}\Big) \\
&= -\frac{\Pi_{x_C,x_C}}{|J|}\Big[-\Big(1-\frac{\partial(l_C r_C)}{\partial x_C}\Big)\frac{(1-\delta)r_C}{r_P} + \frac{\partial x_T}{\partial x_C}\Big(\frac{\partial(l_C r_C)}{\partial x_T}\frac{r_C(1-\delta)}{r_P} + (1-\delta)(1-\frac{\partial l_T}{\partial x_T})\Big)\Big]
\end{aligned}$$

Si $\frac{\partial x_T}{\partial x_C} > 0$, alors $\frac{\partial x_C}{\partial V_U(C)} < 0$ sous la condition (1.33).

Si $\frac{\partial x_T}{\partial x_C} < 0$ il est facile de voir que $\frac{\partial x_C}{\partial V_U(C)} < 0$.

9.8 Preuve du Lemme 1.7

Le cas de l_T. Pour $\theta \leq x_T$ nous avons les relations suivantes entre les distributions :

$$\begin{cases} f(\theta) = \dfrac{q(\theta)(1-\alpha+\alpha l_T)}{1-\alpha+\alpha(x_T-\theta)} & \text{à partir de 1.15} \\ \dfrac{n_T}{N_T}f(\theta) + \dfrac{N_T-n_T}{N_T}q(\theta) = u(\theta) & \text{à partir de 1.27} \end{cases}$$

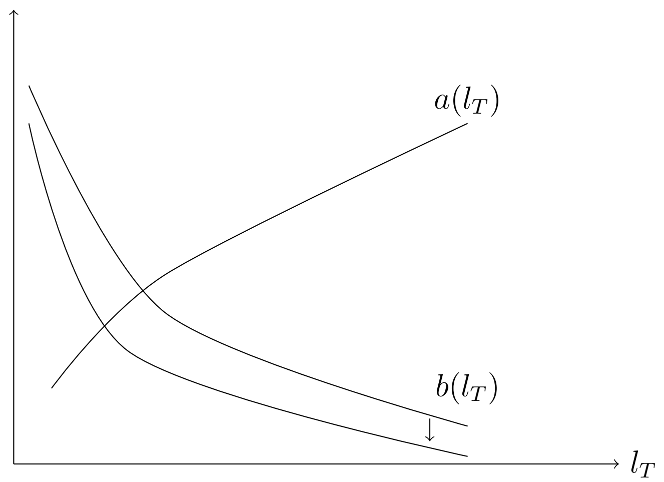
Ainsi, $f(\theta) = \dfrac{u(\theta)N_T(1-\alpha+\alpha l_T)}{N_T(1-\alpha+\alpha(x_T-\theta)) - \alpha n_T(x_T-\theta-l_T)}$

$$\begin{aligned} l_T &= \int_{\underline{\theta}_T}^{x_T} (x_T-\theta)f(\theta)d\theta \\ &= \int_{\underline{\theta}_T}^{x_T} \frac{(x_T-\theta)u(\theta)N_T(1-\alpha+\alpha l_T)}{N_T(1-\alpha+\alpha(x_T-\theta)) - \alpha n_T(x_T-\theta-l_T)} d\theta \end{aligned}$$

Cela se réécrit :

$$\underbrace{\frac{l_T}{(1-\alpha+\alpha l_T)}}_{\equiv a(l_T)} = \underbrace{\int_0^{x_T} \frac{(x_T-\theta)u(\theta)N_T}{N_T(1-\alpha+\alpha(x_T-\theta)) - \alpha n_T(x_T-\theta-l_T)}}_{\equiv b(l_T)} d\theta$$

L'expression à gauche de l'égalité ($a(l_T)$) est croissante par rapport à l_T, alors que l'expression à droite de l'égalité ($b(l_T)$) est décroissante par rapport à l_T. Graphiquement cela peut être représenté de la façon suivante :



Preuve que $b(l_T)$ décroît avec N_T à l_T donné :

$$\begin{aligned} \frac{\partial b(l_T)}{\partial N_T} &= -\int_{\underline{\theta}_T}^{x_T} \frac{\alpha(x_T-\theta)u(\theta)n_T(x_T-\theta-l_T)}{(N_T(1-\alpha+\alpha(x_T-\theta)) - \alpha n_T(x_T-\theta-l_T))^2} d\theta \\ &= -\frac{n_T}{1-\alpha+\alpha l_T} \int_{\underline{\theta}_T}^{x_T} \frac{\alpha(x_T-\theta)f(\theta)(x_T-\theta-l_T)}{(N_T(1-\alpha+\alpha(x_T-\theta)) - \alpha n_T(x_T-\theta-l_T))} d\theta \end{aligned}$$

Pour trouver le signe de $\dfrac{\partial b(l_T)}{\partial N_T}$, nous devons signer :

$$\int_{\underline{\theta}_T}^{x_T} \frac{\alpha(x_T-\theta)f(\theta)(x_T-\theta-l_T)}{(N_T(1-\alpha+\alpha(x_T-\theta))-\alpha n_T(x_T-\theta-l_T))} d\theta \equiv M$$

Le terme $\dfrac{(x_T-\theta)}{(N_T(1-\alpha+\alpha(x_T-\theta))-\alpha n_T(x_T-\theta-l_T))}$ est décroissant par rapport à θ. Donc, les termes positifs sont plus fortement pondérés. Ainsi, si l'expression est positive lorsque tous les termes ont les mêmes poids, elle le sera également dans le cas où les termes positifs sont plus fortement pondérés. Dans le cas où tous les termes ont les mêmes poids, nous avons :

$$\alpha \int_{0\underline{\theta}_T}^{x_T} f(\theta)(x_T-\theta-l_T) = k\alpha(l_T - l_T F(x_T)) \geq 0 \Rightarrow M > 0$$

Dans la mesure où $M > 0$ et $\dfrac{\partial b(l_T)}{\partial N_T} = -\dfrac{n_T}{1-\alpha+\alpha l_T} M \Rightarrow \dfrac{\partial b(l_T)}{\partial N_T} < 0$

A partir du graphique ci-dessus il est immédiat de voir que l_T décroît avec N_T. De façon plus générale, l_T diminue lorsque le taux de chômage augmente (N_T plus élevé ou n_T plus faible).

L'effet sur $(l_C r_C)$**.**

$$\begin{cases} \varphi(\theta) = \dfrac{p(\theta)((1-\alpha)r_C+\alpha(l_C r_C))}{(1-\alpha)r_C+\alpha(x_C-\theta)} & \text{à partir de 1.16} \\ \dfrac{n_C}{N_C}\varphi(\theta) + \dfrac{N_C-n_C}{N_C}p(\theta) = s(\theta) & \text{à partir de 1.28} \end{cases}$$

Ainsi, nous avons :

$$\begin{aligned} (l_C r_C) &= \int_{\underline{\theta}}^{x_C} (x_C-\theta)\varphi(\theta)d\theta \\ &= \int_{\underline{\theta}}^{x_C} \frac{(x_C-\theta)s(\theta)N_C((1-\alpha)r_C+\alpha(l_C r_C))}{N_C((1-\alpha)r_C+\alpha(x_C-\theta))-\alpha n_C(x_C-\theta-(l_C r_C))} d\theta \end{aligned}$$

Cela se réécrit :

$$\underbrace{\frac{(l_C r_C)}{(1-\alpha)r_C + \alpha(l_C r_C)}}_{\eta(l_C r_C)} = \underbrace{\int_{\underline{\theta}_C}^{x_C} \frac{(x_C-\theta)s(\theta)N_C}{N_C((1-\alpha)r_C + \alpha(x_C-\theta)) - \alpha\frac{r_C}{1-x_C+(l_C r_C)}(x_C-\theta-(l_C r_C))}d\theta}_{\equiv\beta(l_C r_C)}$$

Il est facile de voir que $\eta(l_C r_C)$ est croissant par rapport à $(l_C r_C)$. Comme précédemment $\beta(l_C r_C)$ décroît avec $(l_C r_C)$. En effet :

$$\begin{aligned}\frac{\partial\beta(l_C r_C)}{\partial(l_C r_C)} &= -\int_{\underline{\theta}_C}^{x_C} \frac{(x_C-\theta)s(\theta)N_C r_C \alpha(1-\theta)}{(N_C(1-x_C+(l_C r_C))((1-\alpha)r_C+\alpha(x_C-\theta))-\alpha r_C(x_C-\theta-(l_C r_C)))^2}d\theta \\ &< 0\end{aligned}$$

Ainsi, $\dfrac{\partial(l_C r_C)}{\partial N_C} < 0$ peut être démontré en utilisant le même raisonnement que précédemment.

9.9 Preuve de la Proposition 1.2

La variation de z_T a les effets suivants sur les seuils de performance : $\dfrac{\partial x_T}{\partial z_T} < 0$ and $\dfrac{\partial x_C}{\partial z_T} < 0$ (voir la Proposition 1.1).

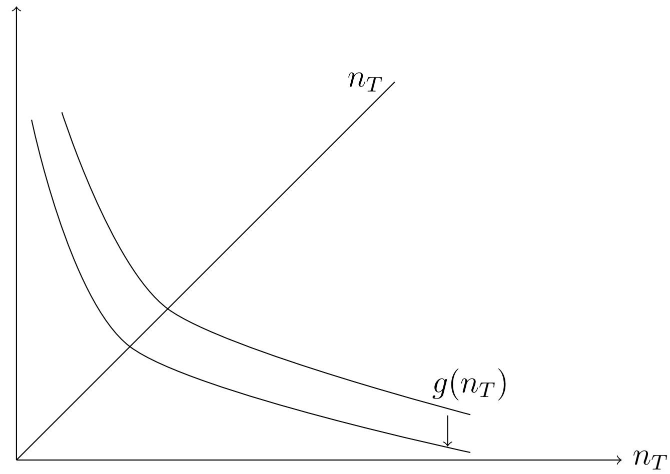
Preuve que $\dfrac{\partial n_T}{\partial z_T} < 0$

$$n_T = \frac{1}{1-x_C+l_C r_C}$$

avec

$$\begin{aligned}(l_C r_C) &= \int_{\underline{\theta}_C}^{x_C} \frac{(x_C-\theta)s(\theta)N_C((1-\alpha)r_C+\alpha(l_C r_C))}{N_C((1-\alpha)r_C+\alpha(x_C-\theta))-\alpha n_C(x_C-\theta-(l_C r_C))}d\theta \\ &= \int_{\underline{\theta}_C}^{x_C} \frac{(x_C-\theta)s(\theta)N_C((1-\alpha)r_C+\alpha(l_C r_C))}{N_C((1-\alpha)r_C+\alpha(x_C-\theta))-\alpha n_T r_C(x_C-\theta-(l_C r_C))}d\theta\end{aligned}$$

En appliquant le résultat obtenu en Annexe 9.8, nous pouvons dire que $(l_C r_C)$ augmente avec n_T. Donc, $\dfrac{1}{1-x_C+l_C r_C} \equiv g(n_T)$ diminue avec n_T.



Si $\dfrac{\partial g(n_T)}{\partial z_T} < 0$ à n_T donné, alors $\dfrac{\partial n_T}{\partial z_T} < 0$.

$$\frac{\partial g(n_T)}{\partial z_T} = \frac{1}{(1-x_C+l_Cr_C)^2}\Big(\underbrace{(1-\frac{\partial(l_Cr_C)}{\partial x_C})\frac{\partial x_C}{\partial z_T}}_{-} - \underbrace{\frac{\partial(l_Cr_C)}{\partial x_T}}_{-}\underbrace{\frac{\partial x_T}{\partial z_T}}_{-}\Big)$$

Pour montrer que $\dfrac{\partial(l_Cr_C)}{\partial x_T} < 0$, nous appliquons les résultats des Annexes 9.4 et 9.8.

Preuve que $\dfrac{\partial n_C}{\partial z_T} \gtrless 0$

$$n_C = \frac{1-x_T+l_T}{1-x_C+l_Cr_C}$$

Nous savons que $\dfrac{\partial(l_Cr_C)}{\partial n_C} > 0$, donc $\dfrac{1-x_T+l_T}{1-x_C+l_Cr_C} \equiv d(n_C)$ est décroissant par rapport à n_C.

Si $\dfrac{\partial d(n_C)}{\partial z_T} \gtrless 0$ à n_C donné, alors $\dfrac{\partial n_C}{\partial z_T} \gtrless 0$.

$$\frac{\partial d(n_C)}{\partial z_T} = \frac{1}{r_P^2}\Big(\underbrace{-(1-\frac{\partial l_T}{\partial x_T})r_P\frac{\partial x_T}{\partial z_T}}_{+}\underbrace{+(1-\frac{\partial(l_Cr_C)}{\partial x_C})r_C\frac{\partial x_C}{\partial z_T}}_{-}\underbrace{-\frac{\partial(l_Cr_C)}{\partial x_T}\frac{\partial x_T}{\partial z_T}r_C}_{-}\Big) \gtrless 0$$

Dans les simulations, les deux situations peuvent apparaître.

Preuve que $\dfrac{\partial s_C}{\partial z_T} \gtreqless 0$

$$s_C = \frac{c(1-\alpha\delta)}{\alpha\delta} + c\frac{x_C - x_T + l_T}{1 - x_T + l_T} + z_C + \frac{c\lambda_C}{\alpha}$$

$$\frac{\partial s_C}{\partial z_T} = \underbrace{\frac{c}{r_C}\frac{\partial x_C}{\partial z_T}}_{-} \underbrace{- \frac{c(1-x_C)}{r_C^2}\frac{\partial x_T}{\partial z_T}}_{+} + \frac{c}{\alpha}\frac{\partial \lambda_C}{\partial z_T}$$

Dans les simulations, le premier effet l'emporte systématiquement, c'est-à-dire que nous avons : $\dfrac{\partial s_C}{\partial z_T} < 0$.

Ci-dessous nous présentons deux cas de simulations où les effets d'une variation de l'indemnité de chômage des travailleurs (z_T) sur le nombre de cadre (n_C) ne sont pas les mêmes. Dans les deux cas, le salaire des cadres décroît avec z_T.

Nous prenons les valeurs suivantes pour les paramètres :

$\alpha = 0.9$ $\quad \delta = 0.95$ $\quad \underline{\theta}_T = 0$ $\quad \overline{\theta}_T = 0.4$ $\quad \underline{\theta}_C = 0.4$ $\quad \overline{\theta}_C = 0.8$

$\theta^P = 1$ $\quad z_C = 0.05$ $\quad N_T = 3$ $\quad N_C = 2.5$

c	0.1						0.05				
z_T	0.02	0.03	0.04	0.05	0.06		0.02	0.03	0.04	0.05	0.06
x_T	0.3020	0.3018	0.3016	0.3014	0.3012		0.3211	0.3209	0.3207	0.3205	0.3204
x_C	0.7484	0.7482	0.7480	0.7477	0.7475		0.7558	0.7558	0.7557	0.7556	0.7556
l_T	0.1071	0.1070	0.1068	0.1067	0.1065		0.1218	0.1217	0.1216	0.1214	0.1213
$l_C r_C$	0.1422	0.1420	0.1418	0.1417	0.1415		0.1486	0.1486	0.1485	0.1485	0.1484
n_T	2.5391	2.5390	2.5388	2.5386	2.5384		2.5457	2.5456	2.5456	2.5455	2.5454
n_C	2.0444	2.0444	2.0444	2.0444	2.0444		2.0385	2.0385	2.0386	2.0386	2.0387
s_T	0.1874	0.1973	0.2071	0.2169	0.2268		0.1098	0.1197	0.1297	0.1396	0.1496
s_C	0.2648	0.2647	0.2646	0.2644	0.2643		0.1588	0.1588	0.1587	0.1587	0.1587
Π	1.5219	1.4970	1.4721	1.4473	1.4225		1.9425	1.9172	1.8918	1.8665	1.8412

9.10 Preuve de la Proposition 1.3

La variation de z_C a les effets suivants sur les seuils de performance : $\dfrac{\partial x_T}{\partial z_C} > 0$ et $\dfrac{\partial x_C}{\partial z_C} < 0$ (voir la Proposition 1.1).

Preuve que $\dfrac{\partial n_T}{\partial z_C} \lesseqgtr 0$

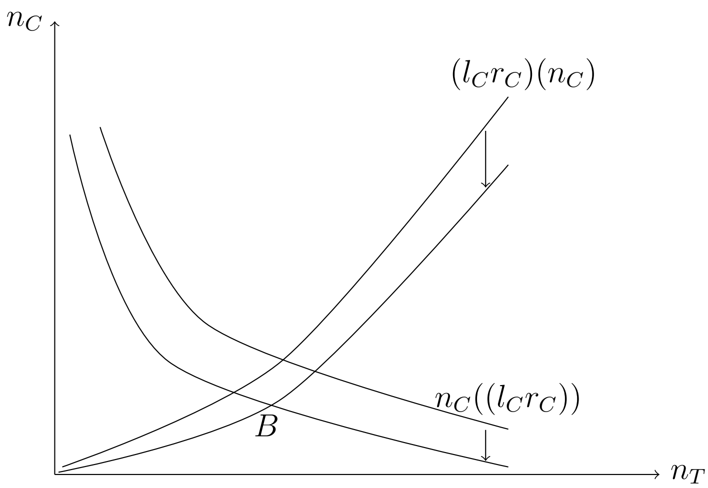
Si $\dfrac{\partial g(n_T)}{\partial z_C} \lesseqgtr 0$ à n_T donné, alors $\dfrac{\partial n_T}{\partial z_C} \lesseqgtr 0$.

$$\frac{\partial g(n_T)}{\partial z_C} = \frac{1}{(1 - x_C + (l_C r_C))^2}\Big(\underbrace{(1 - \frac{\partial(l_C r_C)}{\partial x_C})\frac{\partial x_C}{\partial z_C}}_{-} - \underbrace{\frac{\partial(l_C r_C)}{\partial x_T}}_{-}\underbrace{\frac{\partial x_T}{\partial z_C}}_{+}\Big)$$

Il y a deux effets qui jouent en sens inverses. Dans les simulations, les deux situations apparaissent.

Preuve que $\dfrac{\partial n_C}{\partial z_C} < 0$

$n_C = \dfrac{1 - x_T + l_T}{1 - x_C + (l_C r_C)}$, donc n_C décroît avec $(l_C r_C)$. De l'autre côté $(l_C r_C)$ augmente avec n_C.



A $(l_C r_C)$ donné, $\dfrac{\partial n_C}{\partial z_C} = -\dfrac{(1 - \frac{\partial l_T}{\partial x_T})}{r_P}\dfrac{\partial x_T}{\partial z_C} + \dfrac{r_C}{r_P^2}\dfrac{\partial x_C}{\partial z_C} < 0.$

A n_C donné, $\dfrac{d(l_C r_C)}{dz_C} = \dfrac{\partial(l_C r_C)}{\partial x_C}\dfrac{\partial x_C}{\partial z_C} + \dfrac{\partial(l_C r_C)}{\partial x_T}\dfrac{\partial x_T}{\partial z_C} < 0$. Les nouvelles valeurs d'équilibre pour n_T et n_C correspondent au point B : n_C décroît.

Ci-dessous nous présentons deux cas de simulations où le nombre de travailleurs (n_T) peut respectivement augmenter ou diminuer suite à un accroissement de l'indemnité de chômage des cadres z_C.

Nous prenons les valeurs suivantes pour les paramètres :

$$\alpha = 0.9 \qquad \delta = 0.95 \qquad \underline{\theta}_T = 0 \qquad \overline{\theta}_T = 0.4 \qquad \underline{\theta}_C = 0.4 \qquad \overline{\theta}_C = 0.8$$
$$\theta^P = 1 \qquad z_T = 0.02 \qquad N_T = 3 \qquad N_C = 2.5$$

c	0.05						0.15				
z_C	0.05	0.06	0.07	0.08	0.09		0.05	0.06	0.07	0.08	0.09
x_T	0.3211	0.3229	0.3245	0.3260	0.3274		0.2898	0.2914	0.2930	0.2945	0.2959
x_C	0.7558	0.7558	0.7557	0.7556	0.7556		0.7356	0.7351	0.7346	0.7341	0.7336
l_T	0.1218	0.1233	0.1246	0.1258	0.1269		0.0982	0.0993	0.1004	0.1015	0.1025
$l_C r_C$	0.1486	0.1486	0.1485	0.1484	0.1484		0.1313	0.1309	0.1304	0.1300	0.1295
n_T	2.5457	2.5457	2.5458	2.5459	2.5460		2.5272	2.5269	2.5266	2.5263	2.5259
n_C	2.0385	2.0376	2.0369	2.0362	2.0355		2.0430	2.0416	2.0401	2.0388	2.0374
s_T	0.1098	0.1103	0.1108	0.1112	0.1116		0.2567	0.2578	0.2588	0.2597	0.2606
s_C	0.1588	0.1686	0.1785	0.1883	0.1982		0.3598	0.3687	0.3776	0.3865	0.3954
Π	1.9425	1.9214	1.9004	1.8794	1.8584		1.1433	1.1228	1.1025	1.0823	1.0622

9.11 Preuve de la Proposition 1.5

La variation θ^P a les effets suivants sur les seuils d'exigence : $\dfrac{\partial x_T}{\theta^P} > 0$ et $\dfrac{\partial x_C}{\theta^P} > 0$ (voir la Proposition 1.1).

Preuve que $\dfrac{\partial n_T}{\partial \theta^P} > 0$
Si $\dfrac{\partial g(n_T)}{\partial \theta^P} \lesseqgtr 0$ à n_T donné, alors $\dfrac{\partial n_T}{\partial \theta^P} > 0$.

$$\frac{\partial g(n_T)}{\partial \theta^P} = \frac{1}{(1 - x_C + (l_C r_C))^2}\Big(\underbrace{(1 - \frac{\partial (l_C r_C)}{\partial x_C})\frac{\partial x_C}{\partial \theta^P}}_{+} - \underbrace{\frac{\partial (l_C r_C)}{\partial x_T}}_{-}\underbrace{\frac{\partial x_T}{\partial \theta^P}}_{+}\Big) > 0 \Rightarrow \frac{\partial n_T}{\partial \theta^P} > 0$$

Preuve que $\dfrac{\partial n_C}{\partial \theta^P} > 0$
Si $\dfrac{\partial d(n_C)}{\partial \theta^P} < 0$ à n_C donné, alors $\dfrac{\partial n_C}{\partial \theta^P} < 0$.

$$\frac{\partial d(n_C)}{\partial \theta^P} = \frac{1}{r_P^2}\Big(\underbrace{-(1 - \frac{\partial l_T}{\partial x_T})r_P\frac{\partial x_T}{\partial \theta^P}}_{-}\underbrace{+(1 - \frac{\partial (l_C r_C)}{\partial x_C})r_C\frac{\partial x_C}{\partial \theta^P}}_{+}\underbrace{-\frac{\partial (l_C r_C)}{\partial x_T}\frac{\partial x_T}{\partial \theta^P}r_C}_{+}\Big) \gtreqless 0$$

Preuve que $\dfrac{\partial s_T}{\partial \theta^P} > 0$

$$s_T = \frac{c(1-\alpha\delta)}{\alpha\delta} + c x_T + z_T + \frac{c\lambda_T}{\alpha}$$
$$\frac{\partial s_T}{\partial \theta^P} = c\Big(\underbrace{\frac{\partial x_T}{\partial \theta^P}}_{+} + \underbrace{\frac{1}{\alpha}\frac{\partial \lambda_T}{\partial \theta^P}}_{+}\Big) > 0$$

λ_T augmente avec θ^P puisqu'à la fois n_T et l_T augmentent avec θ^P.

Preuve que $\dfrac{\partial s_C}{\partial \theta^P} \gtreqless 0$

$$s_C = \frac{c(1-\alpha\delta)}{\alpha\delta} + c\frac{(x_C - x_T + l_T)}{1 - x_T + l_T} + z_C + \frac{c\lambda_C}{\alpha}$$

$$\frac{\partial s_C}{\partial \theta^P} = \underbrace{\frac{c}{r_C}\frac{\partial x_C}{\partial \theta^P}}_{+}\underbrace{-\frac{c(1-x_C)}{r_C^2}\frac{\partial x_T}{\partial \theta^P}}_{-} + \frac{c}{\alpha}\frac{\partial \lambda_C}{\partial z_T}$$

Dans les simulations, le premier effet l'emporte systématiquement, c'est-à-dire que nous avons : $\dfrac{\partial s_C}{\partial \theta^P} > 0$.

Ci-dessous nous présentons deux cas de simulations où le nombre de cadres (n_C) respectivement augmente et diminue suite à un accroissement de la productivité du principal (θ^P). Nous remarquons que, dans les deux cas, le salaire des cadres augmente avec θ^P.

Nous prenons les valeurs suivantes pour les paramètres :

$$\alpha = 0.9 \quad \delta = 0.95 \quad \underline{\theta}_T = 0 \quad \overline{\theta}_T = 0.4 \quad \underline{\theta}_C = 0.4 \quad \overline{\theta}_C = 0.8$$
$$z_T = 0.04 \quad z_C = 0.08 \quad N_T = 3 \quad N_C = 2.5$$

c	0.15						0.05				
θ^P	0.8	0.85	0.9	0.95	1		0.8	0.85	0.9	0.95	1
x_T	0.2873	0.2894	0.2912	0.2927	0.2940		0.3227	0.3235	0.3243	0.3250	0.3257
x_C	0.7117	0.7190	0.7248	0.7293	0.7329		0.7536	0.7542	0.7547	0.7551	0.7555
l_T	0.0961	0.0977	0.0990	0.1002	0.1012		0.1231	0.1238	0.1244	0.1250	0.1256
$l_C r_C$	0.1118	0.1176	0.1222	0.1259	0.1289		0.1466	0.1472	0.1476	0.1480	0.1483
n_T	2.4995	2.5092	2.5161	2.5212	2.5250		2.5444	2.5448	2.5452	2.5455	2.5458
n_C	2.0215	2.0280	2.0326	2.0358	2.0380		2.0366	2.0365	2.0364	2.0363	2.0362
s_T	0.2638	0.2690	0.2730	0.2762	0.2788		0.1300	0.1303	0.1306	0.1308	0.1311
s_C	0.3600	0.3687	0.3755	0.3808	0.3850		0.1877	0.1879	0.1881	0.1882	0.1883
Π	1.1125	1.0866	1.0659	1.0495	1.0363		1.8312	1.8305	1.8298	1.8292	1.8286

9.12 Simulations par rapport à N_T

Ici, nous présentons quelques examples de simulations par rapport à la taille de la population des agents non-diplômés (N_T).

$$\alpha = 0.9 \quad \delta = 0.95 \quad \underline{\theta}_T = 0 \quad \overline{\theta}_T = 0.4 \quad \underline{\theta}_C = 0.4 \quad \overline{\theta}_C = 0.8$$
$$\theta^P = 0.9 \quad z_T = 0.02 \quad z_C = 0.05 \quad N_C = 2.5$$

c	0.05						0.1				
N_T	2.8	2.85	2.9	2.95	3		2.8	2.85	2.9	2.95	3
x_T	0.3162	0.3171	0.3179	0.3186	0.3193		0.2955	0.2969	0.2981	0.2990	0.2999
x_C	0.7547	0.7548	0.7550	0.7550	0.7551		0.7427	0.7440	0.7449	0.7455	0.7459
l_T	0.1208	0.1208	0.1207	0.1206	0.1204		0.1050	0.1053	0.1055	0.1056	0.1056
$l_C r_C$	0.1478	0.1479	0.1480	0.1480	0.1480		0.1375	0.1386	0.1393	0.1397	0.1401
n_T	2.5436	2.5441	2.5444	2.5448	2.5451		2.5332	2.5347	2.5357	2.5365	2.5371
n_C	2.0466	2.0446	2.0426	2.0408	2.0390		2.0506	2.0491	2.0474	2.0458	2.0441
s_T	0.1593	0.1407	0.1273	0.1171	0.1092		0.2717	0.2406	0.2175	0.1998	0.1857
s_C	0.1598	0.1595	0.1592	0.1589	0.1586		0.2634	0.2637	0.2637	0.2635	0.2632
Π	1.8114	1.8599	1.8953	1.9224	1.9437		1.3048	1.3844	1.4442	1.4907	1.5279

$\alpha = 0.9 \quad \delta = 0.95 \quad \underline{\theta}_T = 0 \quad \overline{\theta}_T = 0.3 \quad \underline{\theta}_C = 0.5 \quad \overline{\theta}_C = 0.7$

$\theta^P = 0.9 \quad z_T = 0.02 \quad z_C = 0.05 \quad N_C = 2.5$

c	0.05						0.1				
N_T	2.8	2.85	2.9	2.95	3		2.8	2.85	2.9	2.95	3
x_T	0.2320	0.2325	0.2329	0.2333	0.2337		0.2181	0.2187	0.2192	0.2197	0.2202
x_C	0.6850	0.6851	0.6851	0.6852	0.6852		0.6785	0.6790	0.6793	0.6796	0.6798
l_T	0.0867	0.0866	0.0865	0.0864	0.0863		0.0764	0.0764	0.0764	0.0763	0.0763
$l_C r_C$	0.0830	0.0831	0.0831	0.0831	0.0831		0.0771	0.0776	0.0779	0.0781	0.0783
n_T	2.5125	2.5126	2.5127	2.5128	2.5129		2.5085	2.5089	2.5092	2.5094	2.5096
n_C	2.1475	2.1461	2.1448	2.1436	2.1424		2.1531	2.1519	2.1507	2.1495	2.1484
s_T	0.1265	0.1137	0.1042	0.0969	0.0911		0.2201	0.1967	0.1793	0.1657	0.1549
s_C	0.1535	0.1532	0.1530	0.1527	0.1525		0.2542	0.2541	0.2539	0.2536	0.2533
Π	1.8650	1.8980	1.9227	1.9419	1.9573		1.4090	1.4684	1.5133	1.5485	1.5768

9.13 Simulations par rapport à N_C

Ici nous présentons quelques exemples de simulations par rapport à la taille de la population des agents diplômés (N_C).

$\alpha = 0.9$ $\delta = 0.95$ $\underline{\theta}_T = 0$ $\overline{\theta}_T = 0.4$ $\underline{\theta}_C = 0.4$ $\overline{\theta}_C = 0.8$
$\theta^P = 0.9$ $z_T = 0.04$ $z_C = 0.08$ $N_T = 2.8$

c	0.05						0.12				
N_C	2.2	2.25	2.3	2.35	2.4		2.2	2.25	2.3	2.35	2.4
x_T	0.3333	0.3303	0.3279	0.3259	0.3242		0.3151	0.3109	0.3070	0.3034	0.3001
x_C	0.7509	0.7517	0.7523	0.7529	0.7533		0.7131	0.7187	0.7226	0.7255	0.7277
l_T	0.1340	0.1316	0.1298	0.1282	0.1269		0.1187	0.1157	0.1129	0.1103	0.1080
$l_C r_C$	0.1495	0.1492	0.1489	0.1485	0.1481		0.1176	0.1213	0.1236	0.1252	0.1263
n_T	2.5091	2.5154	2.5214	2.5273	2.5329		2.4718	2.4838	2.4936	2.5018	2.5092
n_C	2.0091	2.0157	2.0219	2.0277	2.0333		1.9865	1.9990	2.0095	2.0188	2.0271
s_T	0.1708	0.1723	0.1739	0.1757	0.1776		0.3059	0.3115	0.3159	0.3197	0.3232
s_C	0.2796	0.2509	0.2309	0.2163	0.2051		0.4650	0.4285	0.3986	0.3743	0.3543
Π	1.5188	1.5764	1.6161	1.6448	1.6661		0.7920	0.8535	0.9047	0.9463	0.9799

$\alpha = 0.9$ $\delta = 0.95$ $\underline{\theta}_T = 0$ $\overline{\theta}_T = 0.3$ $\underline{\theta}_C = 0.5$ $\overline{\theta}_C = 0.7$
$\theta^P = 0.9$ $z_T = 0.04$ $z_C = 0.08$ $N_T = 2.8$

c	0.05						0.1				
N_S	2.2	2.25	2.3	2.35	2.4		2.2	2.25	2.3	2.35	2.4
x_T	0.2593	0.2532	0.2486	0.2452	0.2424		0.2396	0.2346	0.2303	0.2267	0.2235
x_C	0.6820	0.6833	0.6838	0.6841	0.6844		0.6723	0.6745	0.6758	0.6767	0.6773
l_T	0.1088	0.1036	0.0998	0.0970	0.0948		0.0924	0.0886	0.0853	0.0826	0.0803
$l_C r_C$	0.0822	0.0830	0.0831	0.0831	0.0830		0.0729	0.0745	0.0753	0.0758	0.0760
n_T	2.4989	2.5018	2.5042	2.5064	2.5085		2.4964	2.5001	2.5031	2.5056	2.5079
n_C	2.1227	2.1275	2.1315	2.1351	2.1383		2.1290	2.1351	2.1402	2.1447	2.1488
s_T	0.1527	0.1512	0.1502	0.1496	0.1492		0.2483	0.2469	0.2456	0.2445	0.2436
s_C	0.4052	0.3011	0.2520	0.2236	0.2051		0.4034	0.3559	0.3231	0.2994	0.2816
Π	1.2572	1.4830	1.5910	1.6541	1.6956		1.0177	1.1230	1.1968	1.2508	1.2918

9.14 Récapitulatif des résultats de statique comparative

	$z_T \searrow$	$z_C \searrow$	$\theta^P \nearrow$	$N_T \nearrow$	$N_C \nearrow$
x_T	$+$	$-$	$+$	$+$	$-$
x_C	$+$	$+$	$+$	$+$	$+$
n_T	$+$	$+/-$	$+$	$+$	$+$
n_C	$+/-$	$+$	$+/-$	$-$	$+$
$\frac{n_T}{n_C}$	$+$	$-$	$+$	$+$	$-$
s_T	$-$	$-$	$+$	$-$	$+/-$
s_C	$+/-$	$-$	$+/-$	$+/-$	$-$

Dans la première colonne, par exemple, nous présentons les effets d'une baisse de l'indemnité de chômage des travailleurs sur les seuils d'exigence, le nombre d'employés à chaque niveau de la hiérarchie, le nombre de travailleurs par cadre et les salaires. Nous voyons ainsi que les seuils d'exigence et le nombre de travailleurs augmentent, que le salaire des travailleur diminue et que les effets sont ambigus sur le nombre et le salaire des cadres.

10 Notations

c - coût de l'effort

δ - taux d'escompte

$(1-\alpha)$ - taux de rotation exogène

z_T - l'indemnité de chômage d'un agent qui a travaillé comme « travailleur »

z_C - l'indemnité de chômage d'un agent qui a travaillé comme « cadre »

N_T - la taille de la population des non-diplômés

N_C - la taille de la population des diplômés

θ - la compétence d'un agent

$U(\theta)$ - la fonction de répartition initiale des compétences des non-diplômés

$F(\theta)$ - la fonction de répartition des compétences des travailleurs

$Q(\theta)$ - la fonction de répartition des compétences des non-diplômés au chômage

$S(\theta)$ - la fonction de répartition initiale des compétences des diplômés

$\Phi(\theta)$ - la fonction de répartition des compétences des cadres

$P(\theta)$ - la fonction de répartition des compétences des diplômés au chômage

x_T - seuil d'exigence des travailleurs

x_C - seuil d'exigence des cadres

l_T - proportion de travailleurs licenciés

l_C - proportion de cadres licenciés

r_C - proportion de problèmes transmis par les travailleurs et reçus par les cadres

r_P - proportion de problèmes transmis par les cadres et reçus par le principal

n_T - nombre de travailleurs

n - le nombre de travailleurs par cadre

n_C - nombre de cadres

λ_T - probabilité de réembauche d'un non-diplômé

λ_C - probabilité de réembauche d'un diplômé

$V_U(T)$ - utilité inter-temporelle d'un non-diplômé au chômage

$V_U(C)$ - utilité inter-temporelle d'un diplômé au chômage

s_T - salaire des travailleurs

s_C - salaire des cadres

Chapitre 2

Supply of Skilled Labor and Organizational Change

1 Introduction

Over the past decades firms experienced important organizational transformations. Hierarchies became flatter, and decision making more decentralized (Rajan and Wulf (2006)). The latter organizational transformations modified firm's labor demand in favor of skilled employees (Caroli and Van Reenen (2001)). This chapter offers a theory in which increasing supply of skilled labor may be at the origin of such a skilled biased organizational change. Furthermore we propose a novel explanation for the skilled wage premium puzzle. Indeed in the United States the skilled wage premium[1] fell, in the 1970s and rose in the 1980s and 1990s, in spite of increasing supply of skilled labor. Katz and Murphy (1992) show that a key component for explaining increasing wage inequality is a shift in the relative demand in favor of skilled labor. One of the explanations provided for it is skilled biased technological change[2]. We show that skill biased organizational change may be an alternative

[1] See Katz and Murphy (1992) and Goldin and Katz (1999) for more details.
[2] Autor, Katz and Krueger (1998) and Machin and van Reenen (1998) among others.

explanation for the shift in firms' demand in favor of qualified labor.

In the model economy production is organized in knowledge based hierarchies (*à la* Garicano (2000)). Agents on the first layer of the organization receive problems and solve those they are able to. Unsolved tasks are transmitted to the next layer of the organization. Knowledge is embodied in employees, who can either be skilled or unskilled. Incentives are guaranteed through efficiency wage contracts (high fixed wage combined with a threat of dismissal). When unemployment is low, an employee knows that being fired is not a harsh punishment, because it is relatively easy to find a new job. Thus a lower unemployment raises the wage cost of agents from the corresponding population. When a principal hires skilled agents, he assigns them to the layer below him. The decision to hire an additional layer of unskilled workers responds to the following trade-off : the benefit of three layer organization (with unskilled workers) is that skilled agents focus on problems that have not been solved by their unskilled subordinates, which reduces both the number of skilled and their wages ; the cost is the reward of unskilled employees. A higher supply of skilled labor reduces their reward and decreases the benefit from adopting a three layer organization. So increasing supply of skilled labor may lead to flattering of organizational forms. Then the composition of jobs is modified. Tasks that have been performed by unskilled agents are assigned to skilled ones. Such jobs are associated to higher wages in order to compensate the broader work charge of skilled employees. Skilled wage premium increases.

This chapter builds on the work of Garicano (2000), who derives the conditions under which a knowledge based hierarchy is the optimal way for organizing production. Another paper close to ours is Garicano and Rossi-Hansberg (2006). In a setting where knowledge is essential input in production and heterogeneous in skills employees are organized in knowledge based hierarchies, they study the way agents are matched with each other and the equilibrium structure of earning. In both papers incentive considerations are set aside, and the authors put the accent on the impact of evolutions in the cost of acquiring and transmitting information on the structure of organizations and wages. In our work, the principal has to motivate agents to work and wages depend on labor market conditions. In particular our analysis is on the impact of evolutions in the supply of qualified workforce on firms' organization, and *a fortiori* on the composition of employment and wages.

Finally, this chapter is also related to Acemoglu (1999). He explains the skilled wage premium evolution with a qualitative change in the composition of jobs. In his paper the increasing supply of skilled labor may qualitatively modify the compositions of jobs by affecting the firms' decisions to invest or not in a skill demanding technology. In our model the increase of skilled labor supply alters the structure of organizations rather than investment decisions.

The rest of the chapter is organized as follows. Section 2 introduces the model. Section 3 analyzes a principal's choice between two organizations structures, two versus three layer hierarchy. Section 4 presents the market equilibrium and derives our main results. In Section 5, we discuss our assumptions and results. Then we conclude.

2 Model

The framework of the model is very close to that of Chapter 1. The only difference is that in this chapter agents' ability is observable, potential employees are either $\underline{\theta}$ (unskilled) or $\overline{\theta}$ (skilled). This section can be skipped if the reader desires to tackle the analysis of the firm owner's choice between two and three layer organizations (Section 3), directly.

Economy Our economy is composed of a continuum of measure one[3] of homogeneous firm owners (principals) with skill level θ^P, and a continuum of potential

[3] There is no free entry.

employees (agents). Each potential employee belongs to a population characterized by some ability (skill) level, perfectly revealed by a diploma. We assume that agents are skilled with ability $\overline{\theta}$, or unskilled with ability $\underline{\theta}$ ($\theta^P > \overline{\theta} > \underline{\theta}$). N_s (*resp* N_u) is the exogenously given size of skilled (*resp* unskilled) agents population. All parties in the economy are infinitely lived and discount the future at a common rate δ ($\delta \in [0,1]$).

Production In the firm, in order to produce agents[4] have to perform tasks (problems). All problems are equally valuable to solve, and their price is normalized to one. However, some of them are more difficult than others. The complexity x of a problem is *ex ante* unknown, but it is drawn from a commonly known distribution. Dealing with a task requires time, succeeding it requires ability and effort. Each agent in the firm has one unit of time by period and spends it entirely in dealing with one problem. An employee θ can solve a problem, by spending some effort, only if he has the required ability to do so, $x \leq \theta$. When the problem is $x > \theta$, he is able neither to solve it nor to identify its complexity[5]. The effort cost function for an agent θ is :

$$c(x,\theta) = \begin{cases} c & if \quad x \leq \theta \\ +\infty & if \quad x > \theta \end{cases}$$

For the rest of the chapter, we make the following normalizations. First, the distribution of problems is such that agent's skill level θ, corresponds to the proportion of problems he is able to solve ($i.e.$ $x \rightsquigarrow U[0,1]$). Second, a firm owner is able to solve any task ($i.e.$ $\theta^P = 1$) at no cost.

[4] An agent can only produce if he is employed. We rule out the possibility for an agent to be self-employed.

[5] Think, for example, to a customer service. When a complaining client calls with a problem, the employee in charge with him cannot determine if he would be able to solve or not the problem, without spending some time with the client. Furthermore, if he is not competent to find a solution, it would be difficult to evaluate the exact ability level needed for solving it.

Organization A firm owner can work alone[6] or create a larger organization by hiring agents on the labor market. Within the firm communication is possible. An employee who is unable of solving a problem can transmit it (ask help) to another employee. However communication is time consuming for the receiver. The latter spends his time unit dealing with a transmitted task. Hence, production requires skills, which are embodied in individuals, and matching problems with agents who are able to solve them is costly (*i.e.* time consuming). Then production should be organized in knowledge based hierarchies[7]. On the lowest layer there are less able workers, more able agents are assigned to higher levels (managers). Problems only arrive at the workers' level. If a worker knows the solution of a problem he solves it and production is realized. If not, he transmits it to the agents of the next level of the hierarchy. If a manager knows the solution he solves the problem. If not, he transmits it to the next level and so on. Managers only deal with problems that have not been solved by their subordinates.

Incentives The instantaneous utility of an employee exerting effort writes as $(s - c)$. Any unemployed agent receives an exogenously given unemployment benefit z. Each agent is either employed or unemployed at any point in time.

At the beginning of each period a firm owner with vacant occupations and unemployed agents are matched. The principal proposes a contract to the agents. Each of them accepts or refuses. If an agent refuses the contract he returns to the unemployment pool and the vacancy remains unoccupied for the period. If he accepts, he receives a problem and decides to solve (if able) or transmit it. At the end of a period the principal observes the difficulty of each problem, which employee received it and whether it was solved or not. The contract is executed. At the end of

[6] The instantaneous profit of a firm owner that works alone is equal to 1. Indeed, he deals with one task by period, solves it with probability 1 and without effort cost.

[7] Garicano (2000) shows that when matching problems with agents who know their solution is costly the optimal way for organizing production is a knowledge based hierarchy. Productive workers specialize in receiving problems and dealing with easy ones and managers specialize in solving exceptions.

each period the employment relationship may be interrupted for exogenous reasons which arrives with probability $(1-\alpha)$. This exogenous turnover rate is the same whatever agent's ability or position. Exogenous separation causes an employee to enter the unemployment pool.

A principal motivates agents with efficiency wage contracts *à la* Shapiro and Stiglitz (1984), *i.e.* high fixed wage combined with a threat of dismissal of under-performing agents. As shown by MacLeod and Malcomson (1998), this incentive scheme is the optimal one, if performance is observable but unverifiable by a third party and if occupations are in short supply (*i.e.* there is unemployment). So our analysis applies to contexts with unverifiable performance and unemployment.

For a wage to be incentive compatible, the agent's value from working should be larger than his value from shirking.

$$s - c + \alpha\delta V + (1-\alpha)\delta V_U \geq s + \delta V_U$$

$$V - V_U \geq \frac{c}{\alpha\delta} \tag{2.1}$$

where V is the inter-temporal expected utility of an employee who exerts effort and V_U is the inter-temporal expected utility of an unemployed agent. An employee exerts effort only if his rent from keeping the job is sufficiently high.

3 Two vs Three layer hierarchy

For expositional simplicity, we assume that the incentive compatible wage of unskilled workers is below a minimum wage[8] level $\underline{s}$. Hence, in what follows motivation will be an issue only for skilled agents. Hereafter V (*resp* V_U) is the expected utility of an employed (*resp* unemployed) skilled agent.

[8] This assumption is discussed in Section 6 and formally relaxed in the Appendix.

A principal has the choice among three knowledge based hierarchies[9] : a two layer hierarchy with unskilled employees, a two layer hierarchy with skilled employees, and a three layer hierarchy with unskilled workers and skilled managers. In the main part of the chapter we focus our attention on his choice between a three layer organization and a two layer organization with skilled employees. This corresponds to situations in which the supply of skilled labor is "sufficiently" high. In Section 5.3, we turn back to this point and briefly discuss the principal's choice among those three organizational forms.

Each principal is "too small" to have an individual impact on market conditions. Hence, he chooses organizational structure, taking as given skilled agents' outside opportunity (V_U). In this section we first characterize the principal's profit if he chooses a two or three layer hierarchy. Then, we discuss his choice between these organizations, conditionally on the value of V_U.

3.1 Two layer hierarchy

Recall that in a knowledge based hierarchy tasks arrive only on the first layer of the organization, thus the number of problems a firm deals with, corresponds to the number of employees on that level n_1. A skilled employee asks the principal for help with probability $(1-\overline{\theta})$. Thus $n_1(1-\overline{\theta})$ tasks are transmitted to the principal. The firm owner has one unit of time and can deal with one transmitted problem. His time constraint determines the number of agents to be employed $n_1(1-\overline{\theta}) = 1$. It is assumed that it is costly for a principal not to solve a problem arriving in the firm. A possible justification is that when a firm doesn't solve a received problem, it looses its reputation on the product market.

The inter-temporal expected utility of an agent employed in such organization is :

[9] We focus on situations in which it is always preferable for a principal to employ agents rather than to produce alone.

$$V = s - \overline{\theta}c + \alpha\delta V + (1-\alpha)\delta V_U$$

The incentive constraint (2.1) is binding, therefore the corresponding wage is :

$$s = c\overline{\theta} + \frac{c(1-\alpha\delta)}{\alpha\delta} + (1-\delta)V_U$$

So the principal's profit when adopting two layer hierarchy corresponds to :

$$\begin{aligned} \pi &= n_1(1-s) \\ &= \frac{1}{1-\overline{\theta}}(1-s) \end{aligned} \tag{2.2}$$

3.2 Three layer hierarchy

In a three layer organization there are n_1 unskilled workers that receive problems and n_2 managers. Workers transmit to the next level a proportion $(1-\underline{\theta})$ of the problems they receive. The number of managers hired by a principal is such that they use all their time in helping unskilled workers, *i.e.* $n_2 = (1-\underline{\theta})n_1$. Note that n_2 depends both on the number of workers and on their solving capacity. Indeed, more able workers (*i.e.* higher $\underline{\theta}$) solve larger proportion of problems and need fewer managers to help them.

The probability that a problem is solved neither by a worker, nor by a manager is $(1-\overline{\theta})$. The number of problems the organization deals with is still given by the principal's time constraint : $n_1(1-\overline{\theta}) = 1$. Hence in a three layer hierarchy a principal hires $\frac{1}{1-\overline{\theta}}$ unskilled workers and $\frac{1-\underline{\theta}}{1-\overline{\theta}}$ skilled managers.

In this organization skilled agents solve only tasks that have been transmitted by their subordinates. Thus their expected utility depends on their own, and their subordinates' skill level :

$$V = S - c\frac{\overline{\theta}-\underline{\theta}}{1-\underline{\theta}} + \alpha\delta V + (1-\alpha)\delta V_U$$

The corresponding incentive compatible wage is :

$$S = c\frac{\overline{\theta} - \underline{\theta}}{1 - \underline{\theta}} + \frac{c(1 - \alpha\delta)}{\alpha\delta} + (1 - \delta)V_U$$

$$\begin{aligned} \Pi &= n_1(1 - \underline{s}) - n_2 S \\ &= \frac{1}{1 - \overline{\theta}}(1 - \underline{s}) - \frac{1 - \underline{\theta}}{1 - \overline{\theta}} S \end{aligned} \tag{2.3}$$

is the firm-owner's profit. In a three layer hierarchy skilled agents are hired to deal with problems unskilled employees are not able to solve. Thus both the number of skilled and their wage are lower compared to those in a two layer organization.

Remark 2.1. *For given outside option V_U, being a manager in a three layer organization or a worker in a two layer one provides the same utility level to a skilled agent ($V = \frac{c}{\alpha\delta} + V_U$).*

3.3 Principal's choice

We now characterize the principal's choice between a two and a three layer hierarchy.

Lemma 2.1. *There exists a threshold value $\overline{V}_U = \frac{1}{1-\delta}\left(\frac{\underline{s}}{\underline{\theta}} - \frac{c}{\alpha\delta}\right)$ (corresponding to $\Pi = \pi$), such that :*

- *If $V_U < \overline{V}_U$, then a principal adopts a two layer hierarchy.*
- *If $V_U > \overline{V}_U$, then a principal adopts a three layer hierarchy.*

The decision to adopt a three layer organization responds to the following trade-off. Hiring unskilled workers allows skilled agents to focus on difficult problems, which reduces the number of skilled and their wages. The cost for the principal is the reward of unskilled employees. A higher wage cost for skilled agents (*i.e.* higher V_U) increases the benefit from adopting a three layer organization. In the same way a higher wage cost of unskilled labor (here $\underline{s}$), raises $\overline{V}_U$ and makes more valuable the adoption of a two layer hierarchy.

4 Equilibrium

We first compute the equilibrium value of skilled's outside opportunity as function of firms' employment decisions. Then we turn to the determination of the equilibrium structure of firms and skilled's wage level.

4.1 Equilibrium outside option

The key market variable that determines individual firm behavior is the skilled outside opportunity. An unemployed agent receives the unemployment compensation z and at the end of a period, he may find a new job with probability λ. Reemployment opportunities depend on firms' employment decisions. Let β be the proportion of firms organized as three layer hierarchies, $\beta \in [0,1]$. Then the inter-temporal utility of an unemployed skilled agent corresponds to :

$$V_U(\beta) = z + \delta[\lambda(\beta)V + (1-\lambda(\beta))V_U(\beta)] \Leftrightarrow V_U(\beta) = \frac{z + \frac{c\lambda(\beta)}{\alpha}}{1-\delta}$$

where $\lambda(\beta)$ guarantees the equilibrium of flows in and out of the unemployment pool. In steady state the flow into the unemployment pool is $n(\beta)(1-\alpha)$, where $n(\beta)$ corresponds to the number of occupied skilled agents.

$$n(\beta) = \beta\frac{1-\underline{\theta}}{1-\overline{\theta}} + (1-\beta)\frac{1}{1-\overline{\theta}}$$

The flow out is $\lambda(\beta)(N_s - n(\beta))$. In the population of unemployed a proportion $\lambda(\beta)$ of the agents is hired to replace departures. Thus in equilibrium we have :

$$\begin{aligned} \lambda(\beta) &= \frac{(1-\alpha)n(\beta)}{N_s - n(\beta)} \\ &= \frac{(1-\alpha)(1-\beta\underline{\theta})}{N_s(1-\overline{\theta}) - (1-\beta\underline{\theta})} \end{aligned} \tag{2.4}$$

$V_U(1)$ (*resp* $V_U(0)$) is the value of the outside option when all firms in the economy are three (*resp* two) layer hierarchies. Note that $V_U(0) > V_U(1)$. If all firms adopt an organization with three layers, the demand for skilled labor is lower. There are fewer reemployment possibilities, which reduces skilled's outside opportunity.

4.2 Organizations' structure and wages

Three equilibria zones arise. First, if at $\beta = 1$ it is still optimal for a principal to adopt a three layer hierarchy (*i.e.* $V_U(1) \geq \overline{V}_U$), then all firms adopt it. Conversely, if $V_U(0) \leq \overline{V}_U$, then all firms organize production in two layer hierarchies. Finally, as noticed above $V_U(0) > V_U(1)$, so we could be in a situation where $\overline{V}_U \in]V_U(1), V_U(0)[$. In this case, some firms adopt a three layer organization, while others organize production in a two layer hierarchy[10]. The equilibrium value of β is given by $V_U(\beta) = \overline{V}_U$, thus we obtain $\beta = \dfrac{1}{\underline{\theta}}\Big(1 - \dfrac{\gamma}{1+\gamma} N_s(1-\overline{\theta})\Big)$, where $\gamma = \dfrac{\alpha(\overline{V}_U(1-\delta) - z)}{c(1-\alpha)}$. Note that an increase in the supply of skilled labor reduces the proportion of three-layer organizations.

Proposition 2.1. *There are two thresholds, $N_s^\star$ (corresponding to $V_U(1) = \overline{V}_U$) and $N_s^{\star\star}$ (corresponding to $V_U(0) = \overline{V}_U$), such that $N_s^{\star\star} > N_s^\star$:*

- *If $N_s < N_s^\star$ all firms are three layer hierarchies, $\beta = 1$.*
- *If $N_s > N_s^{\star\star}$ all firms are two layer hierarchies, $\beta = 0$.*
- *If $N_s \in [N_s^\star, N_s^{\star\star}]$, some firms are two layer, others three layer hierarchies, $1 > \beta > 0$. And β decreases with N_s.*

An increase in the supply of skilled labor may lead to a shift in the way firms are organized. Indeed, a higher N_s reduces skilled's wage cost and makes more profitable the adoption of flatter organizations, in which only skilled agents are hired. This result seems consistent with empirical observations. Caroli and van Reenen (2001)

[10] Some skilled agents are employed in firms with two layers others in firms with three layers. According to Remark 2.1 a skilled agent is indifferent between those two positions.

for example point out[11] existing complementarities between organizational change and increasing supply of skilled labor. They show that *"the probability of introducing changes in organizations, such as delayering of hierarchies, is depressed by shortages of skilled workers"*.

Let us now turn to the effect of variation in skilled labor supply on wages. The evolution of the average wage of skilled agents with N_s is represented in Figure 2.1. If a variation of N_s occurs in the regions of pure equilibrium ($N_s < N_s^\star$ or $N_s > N_s^{\star\star}$), then the only effect is the traditional one : a higher supply of skilled agents reduces the reemployment probability and depresses the skilled's wages.

Proposition 2.2. *A higher N_s, when $N_s \in [N_s^\star, N_s^{\star\star}]$ raises skilled agents' average wage, and wage inequality.*

The average wage for skilled agents writes :

$$\begin{aligned} \bar{s} &= \beta S + (1-\beta)s \\ &= \frac{c(1-\alpha\delta)}{\alpha\delta} + c\overline{\theta} - \beta\underline{\theta}c\frac{1-\overline{\theta}}{1-\underline{\theta}} + (1-\delta)\overline{V}_U \end{aligned} \tag{2.5}$$

When $N_s \in [N_s^\star, N_s^{\star\star}]$, a higher supply of skilled labor modifies firms' repartition between three and two layer hierarchies, thus decreasing β. There are more skilled agents in the economy and organizational change shifts the demand in favor of skilled labor force. In fine there is no change in skilled's outside opportunity. However the composition of jobs has qualitatively changed. More firms adopt two layer hierarchies, in which skilled are in charge of broader set of tasks and receive higher wages.

[11] Their analysis is based on a panel of British and French firms.

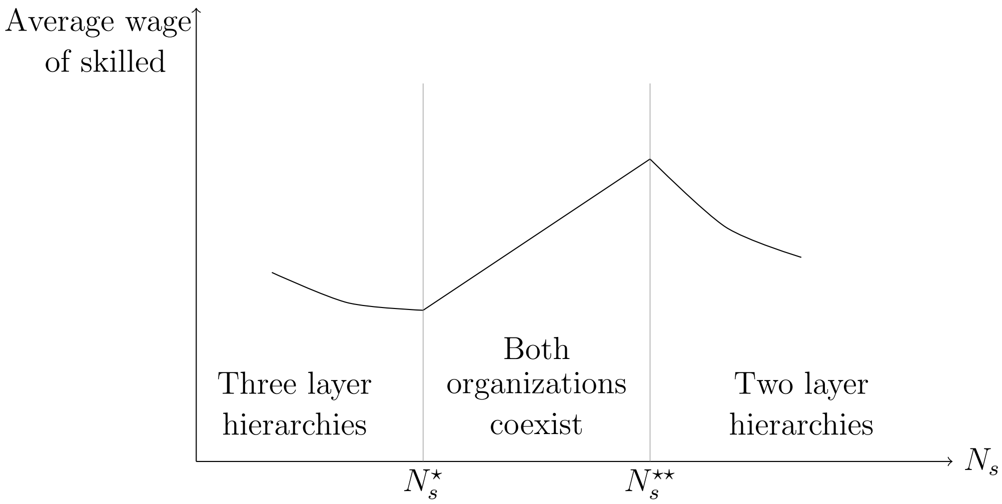


Figure 2.1 : Average wage of skilled agents as function of N_s.

5 Discussion

5.1 Minimum wage evolution

A higher minimum wage, increases $\bar{V}_U$. Thus it reduces the zone in which all firms are three layer hierarchies, and expands the zone in which firms adopt two layer organizations. Inside of the pure equilibrium zones, a higher minimum wage affects neither the employment level of skilled nor of unskilled agents[12]. When the increase of the minimum wage occurs in the zone where both organizational form coexist, a higher $\bar{V}_U$ raises the proportion of firms organized in two layer hierarchies. This adversely affects the demand for unskilled labor, and has a positive effect on skilled's employment and wages.

[12] In a model with free entry, the increase of the minimum wage would lead to the reduction of the number of firms present on the market, thus reducing employment both of skilled and unskilled agents.

5.2 Endogenous wage for unskilled employees

To simplify the exposition we made the assumption of exogenously given wage for unskilled agents. One could wonder at what extend our results depend on it. If we relax the assumption and consider endogenous wage for unskilled employees, the threshold value $\overline{V}_U$ is an increasing function in β. Indeed, larger proportion of three layer hierarchies (higher β) raises the demand for unskilled labor which in turn makes less valuable the adoption of a three layer organization. However, the main findings of our model are robust[13] to the introduction of endogenous wage for unskilled agents. In the zone, where both organizations coexist, a higher supply of skilled labor increases the proportion of flatter organizations. Concerning the average wage of skilled employees, there are two effect. More firms adopt two layer hierarchies in which skilled's reward is higher, thus the average wage increases. At the other side the increase in the demand of skilled labor does not fit the increase in their supply, thus skilled's outside option decreases, and so does the average wage. Even if the positive effect of a higher N_s on skilled's reward is not systematic it still arises.

5.3 Supply of skilled labor and average firm size

We focused our analysis to the comparison of two organizational forms, in order to shed some light on recent evolutions in firms' organization and wage inequalities. In Figure 2.2, we present the effect of evolution in the supply of qualified labor on the average firm size[14] when the principal chooses among the three possible knowledge based organizations.

[13] A formal discussion is proposed in the Appendix.
[14] The size of a firm is measured by the number of employees.

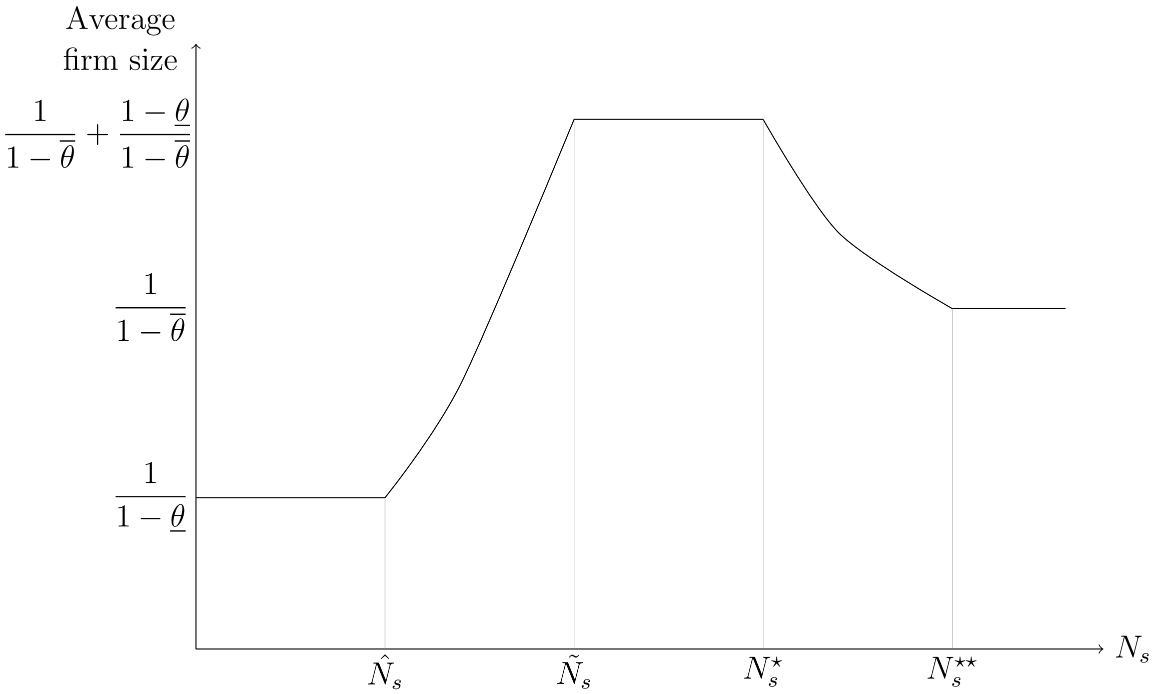


Figure 2.2 : Average firm size as function of N_s.

For a very low supply of skilled labor, here $N_s < \hat{N}_s$, the economy is composed of two layer organizations, in which a firm owner hires unskilled employees only. An increase of N_s leads to the progressive adoption of three layer organizations. This evolution raises the demand both for skilled and unskilled labor. A further increase in skilled's supply leads to delayering. Firms hire fewer but more skillful employees. The evolutions presented on the graph fit the observations documented by Baumol, Blinder and Wolff (2005, pp. 2-4) concerning the size of firms throughout the twentieth century. Indeed, the authors provide evidence for "increasing business size from about 1935 until about 1980, and then decreasing size from about 1980 to about 1993".

6 Conclusion

This chapter offers a model of endogenous choice of firm's organization and job composition. In the model economy, firms are organized in knowledge based hierarchies and employees are assigned to positions conditionally on their skill level. Employees' wages depend on labor market conditions. We discuss the effect of a higher supply of skilled labor on the choice of firm's structure in terms of hierarchical levels and job composition ; and the impact of organizational change on the structure of employment and wage evolution.

The analysis could be extended to situations in which firms are composed of more than two layers, if we consider more than two skill level categories of potential employees. Then, a higher supply of the most skillful would still lead to the adoption of flatter organizations. However there will be situations in which delayering concerns middle managers rather than unskilled workers. The consequence would be a reduction in the proportion of firms employing middle managers, which in turn would reduce the latter's average wage. A more detailed analysis of this point could be a possible extension of the model.

7 References

Acemoglu D. (1999), “Changes in Unemployment and Wage Inequality : An Alternative Theory and some Evidence”, *The American Economic Review*, vol. 89, pp. 1259-1278.

Autor D., L. Katz and A. Krueger (1998), “Computing Inequality : Have Computers Changed the Labor Market ?”, *Quarterly Journal of Economics*, vol. 113, pp. 1169-1213.

Baumol W., A. Blinder and E. Wolff (2005), “Downsizing in America : Reality, Causes, and Consequences”, *Russell Sage Foundation.*

Caroli E. and J. Van Reenen (2001), “Skill Biased Organizational Change ? Evidence from a panel of British and French Establishments”, *Quarterly Journal of Economics*, vol. 116, pp. 1449-1492.

Garicano L. (2000), “Hierarchies and the Organization of Knowledge in Production”, *The Journal of Political Economy*, vol. 108, pp. 874-904.

Garicano L. and E. Rossi-Hansberg (2006), “Organization and Inequality in a Knowledge Economy”, *Quarterly Journal of Economics*, vol. 121, pp. 1383-1436.

Goldin C. and L. Katz (1999), “The Returns to Skill in the United States Across the Century”, *NBER Working Paper* No. 7126.

Katz L. and K. Murphy (1992), “Changes in Relative Wages : Supply and Demand Factors”, *Quarterly Journal of Economics*, vol. 107, pp. 35-78.

Machin S. and J. van Reenen (1998), “Technology and Changes in Skill Structure : Evidence from Seven OECD Countries”, *Quarterly Journal of Economics*, vol. 113, pp. 1215-1244.

MacLeod B. and J. Malcomson (1998), “Motivation and Markets”, *The American Economic Review*, vol. 88, pp. 388-411.

Rajan R. and J. Wulf (2006), "The Flattening Firm : Evidence from Panel Data on the Changing Nature of Corporate Hierarchies", *The Review of Economics and Statistics*, vol. 88, pp. 759-773.

Shapiro C. and J. Stiglitz (1984), "Equilibrium Unemployment as a Worker Discipline Device", *American Economic Review*, vol. 74, pp. 433-444.

8 Appendix

If the wage of unskilled agents is endogenous it corresponds to

$$s_u = \frac{c(1-\alpha\delta)}{\alpha\delta} + c\underline{\theta} + (1-\delta)v_U$$

where v_U is the outside opportunity of unskilled agents, and v is their inter-temporal expected utility when employed. The demand for unskilled labor thus their outside opportunity depends on firms' organization (*i.e.* it is function of β).

$$v_U(\beta) = z_u + \delta(\lambda_u(\beta)v + (1-\lambda_u(\beta))v_U(\beta)) \Leftrightarrow v_U(\beta) = \frac{c\lambda_u(\beta) + \alpha z_u}{\alpha(1-\delta)}$$

$$\begin{aligned}\lambda_u(\beta) &= \frac{(1-\alpha)\frac{\beta}{1-\overline{\theta}}}{N_u - \frac{\beta}{1-\overline{\theta}}} \\ &= \frac{(1-\alpha)\beta}{N_u(1-\overline{\theta}) - \beta}\end{aligned}$$

Thus the threshold value $\bar{V}_U$, also depends on β and writes :

$$\bar{V}_U(\beta) = \frac{c\lambda_u(\beta) + \alpha z_u}{\underline{\theta}\alpha(1-\delta)} + \frac{1}{1-\delta}A \qquad \text{where} \qquad A = \frac{c(1-\alpha\delta)(1-\underline{\theta})}{\alpha\delta\underline{\theta}}$$

The threshold values for N_s change but the results in the zones of pure equilibrium remain the same as with minimum wage. So let us focus to the case of mixed strategy equilibrium. $V_U = \bar{V}_U(\beta)$.

$$\frac{c\lambda(\beta) + z\alpha}{\alpha(1-\delta)} = \frac{c\lambda_u(\beta) + \alpha z_u}{\underline{\theta}\alpha(1-\delta)} + \frac{1}{1-\delta}A \Leftrightarrow \lambda(\beta) = \frac{\lambda_u(\beta)}{\underline{\theta}} + \frac{(\underline{\theta}(A-z) + z_u)\alpha}{c\underline{\theta}}$$

$$\Leftrightarrow \frac{1-\beta\underline{\theta}}{N_s(1-\overline{\theta}) - (1-\beta\underline{\theta})} = \frac{\beta}{\underline{\theta}(N_u(1-\overline{\theta}) - \beta)} + \frac{(\underline{\theta}(A-z) + z_u)\alpha}{\underline{\theta}c(1-\alpha)} \tag{2.6}$$

Result 2.1. *A higher N_s decreases β, i.e. more firms are two layer hierarchies.*

Proof : The derivative of equation (2.6) with respect to N_s writes :

$$-\frac{(1-\beta\underline{\theta})(1-\overline{\theta})}{(N_s(1-\overline{\theta})-(1-\beta\underline{\theta}))^2}-\frac{\partial\beta}{\partial N_s}\frac{(1-\overline{\theta})\underline{\theta}}{(N_s(1-\overline{\theta})-(1-\beta\underline{\theta}))^2}=\frac{\partial\beta}{\partial N_s}\frac{N_u(1-\overline{\theta})}{\underline{\theta}(N_u(1-\overline{\theta})-\beta)^2}$$

$$\Rightarrow \frac{\partial\beta}{\partial N_s}=-\frac{\frac{(1-\beta\underline{\theta})(1-\overline{\theta})}{(N_s(1-\overline{\theta})-(1-\beta\underline{\theta}))^2}}{\frac{(1-\overline{\theta})\underline{\theta}}{(N_s(1-\overline{\theta})-(1-\beta\underline{\theta}))^2}+\frac{N_u(1-\overline{\theta})}{\underline{\theta}(N_u(1-\overline{\theta})-\beta)^2}}<0$$

Result 2.2. *A higher N_s raises the unemployment of skilled agents.*

$\bar{V}_U$ decreases with N_s. Thus a higher N_s leads to larger adoption of flatter (two layer) organizations. However since it decreases unskilled outside opportunities the increase in the demand of skilled labor does not fit the increase in the supply of skilled labor.

There are two effects on skilled average wage and on the wage gap. More firms adopt flatter organizations in which wages are higher, thus the average wage increases. However there are more unemployed skilled agents which decreases their outside opportunity and adversely affects their wages.

Result 2.3. *A higher N_s raises the skilled wage premium for some values of the parameters.*

With endogenous wage for unskilled agents, the increase of the skilled wage premium in the zone of mixed strategy equilibrium is not systematic, however it is still possible.

The average wage for skilled is :

$$\overline{s}=\frac{c(1-\alpha\delta)}{\alpha\delta}+c\overline{\theta}-\beta\underline{\theta}c\frac{(1-\overline{\theta})}{(1-\underline{\theta})}+(1-\delta)\bar{V}_U(\beta)$$

$$\frac{\partial\overline{s}}{\partial N_s}=\frac{\partial\beta}{\partial N_s}\frac{c}{\alpha}\Big(\frac{(1-\alpha)N_u(1-\overline{\theta})}{\underline{\theta}(N_u(1-\overline{\theta})-\beta)^2}-\alpha\underline{\theta}\frac{(1-\overline{\theta})}{(1-\underline{\theta})}\Big)$$

According to Result (1), $\frac{\partial\beta}{\partial N_s}<0$. So the skilled wage increases with N_s only if

$\left(\frac{(1-\alpha)N_u(1-\overline{\theta})}{\underline{\theta}(N_u(1-\overline{\theta})-\beta)^2}-\alpha\underline{\theta}\right) < 0$. For any value of the parameters there exists a threshold value $\bar{\alpha} < 1$, such that the expression above is negative if $\alpha > \bar{\alpha}$.

A similar condition exists for the wage gap, but it is less stringent, since the more extended adoption of flatter organizations reduces unskilled outside opportunities and thus the wage of unskilled employees.

9 Notations

c - effort cost

δ - discount factor

z - the unemployment benefit of skilled unemployed agents

$(1-\alpha)$ - probability of exogenous interruption of the employer-employee relationship

θ - agent's ability, $\underline{\theta}$ for unskilled agents and $\overline{\theta}$ for skilled ones

N_u - the size of unskilled agents population

N_s - the size of skilled agents population

$n(\cdot)$ - the number of employed skilled

s - the wage of a skilled agent employed in two layer hierarchy

S - the wage of a skilled agent employed in three layer hierarchy

$\underline{s}$ - the wage of an unskilled employee

V - the inter-temporal expected utility of an employed skilled

V_U - the inter-temporal expected utility of an unemployed skilled

β - the proportion of firms organized in three layer hierarchies

$\lambda(\cdot)$ - the reemployment probability of an unemployed skilled agent

Chapitre 3

Motivate and Select - Markets and the Form of Compensation

1 Introduction

Performance-based pay is not the only way to motivate employees, as could be suggested by the traditional principal-agent approach. Indeed, in addition of piece rates, employers use a variety of other economic incentives, such as profit sharing, firing, promotion or bonuses based on informal agreements[1]. The use of performance-based pay is limited by the difficulty of measuring individual performance in an objective way. Indeed, in many instances, the available performance measures may be observable for the contracting agents, but too hard to verify by a third party. Yet this information is still useful to motivate employees since informal agreements can be tied on it. Such "relational contracts" are not enforced by a court but by the potential loss of cooperation between the contracting parties. The theory based on "relational incentive contracts" provides a useful understanding of the existence of

[1] MacLeod and Parent (1997) analyze different incentive schemes used in the US. Hayes and Shaefer (1997) provide evidence for the use of subjective performance measures when boards of directors decide the salary and bonus of chief executives. Cappelli and Chauvin (1991) provide evidence in favor of the use of efficiency wages as an instrument to motivate agents to work.

some of the incentive tools mentioned above.

The early literature on relational contracts focuses on frameworks with symmetric information (see Shapiro and Stiglitz (1984) and Bull (1987) among others). MacLeod and Malcomson (1989) propose the most complete treatment of the problem. They prove that provided there is a sufficient rent from employment, either performance-based bonus (workers who perform satisfactory are paid an end of period bonus) or an efficiency wage (high fixed wage combined with a threat of dismissal) could be sustained in equilibrium. The rent corresponds to the surplus[2] generated by the continuation of the employment relationship. If the latter is interrupted each party receives an exogenously given outside opportunity. MacLeod and Malcomson (1998) go further and stress the importance of labor market conditions on the choice between efficiency wage and bonus. When agents are homogeneous in skills and there is unemployment, there is no rent for a principal to continue the employment relationship. The principal has the possibility to replace an employee with an unemployed agent at no cost, since agents are substitutes. Then, the efficiency wage is the only way to motivate employees. Conversely, when employees are in short supply, vacancies cannot be immediately filled, which creates a rent from keeping an agent. Then a performance-based bonus is credible. MacLeod and Malcomson (1998) show if the cost of unoccupied vacations is not too high, then an equilibrium with full employment and bonus could emerge.

However, in real world contracts different incentive tools may coexist[3]. Nevertheless, there are few theoretical explanations[4] for this coexistence. In this paper we propose a rationale for the simultaneous use of termination contract (fixed wage and minimum performance standard to be achieved in order to keep occupation) and bonus when information about employee's production is unverifiable. We

[2] Since both the principal and employees are risk neutral, the reward is a pure transfer.

[3] See for example Baker, Gibbs and Holmström (1994) for a study of the wage policy of a firm.

[4] Baker, Gibbons and Murphy (1994) show that when a principal has available both verifiable and unverifiable performance measure, it can be in his interest to use both measures thus proposing a formal performance-based pay and an informal bonus.

examine this issue in a setting with heterogeneous in unobservable ability agents and endogenous, depending on labor market conditions, outside opportunities.

Levin (2003) characterizes the optimal relational contract with hidden information. Since the information asymmetry parameter changes from period to period, his focus is on the restrictions on revelation due to the self-enforcing character of the agreement. In our model, agents are privately informed about their ability which is the same over time. Under type persistence, revelation is a costly strategy for the principal. Indeed, when performance is not verifiable a principal can end the contract in any period. Then an agent is reluctant to reveal his information since it allows the employer to reap his surplus. In the model, we explicitly set aside *ex ante* screening and focus on the effects of *ex post* selection[5] on the optimal relational contract.

We build our analysis in a framework with a continuum of firm owners proposing homogeneous positions. A job corresponds to the set of tasks an employee is supposed to perform[6]. The task an employee receives in a given period may be more or less difficult. Since the principal observes the difficulty at the end of the period, succeeding a task is a signal about the agent's ability. The firm owner's problem is to motivate agents to perform the largest possible set of tasks. If employees are in short supply the equilibrium contract is still a performance-based bonus. Hence we focus on situations with unemployment. To give a flavor of the analysis consider that employees have larger average ability than unemployed[7].

When agents are heterogeneous, there could be a rent for the principal from keeping a successful insider. Indeed, solving a hard task signals high productivity, that may exceed the productivity of the average unemployed agent. Since the gap

[5] At the end of a period a principal only keeps agents that have attained the performance standard set in the termination contract.

[6] Our production functions is inspired by the one presented in Garicano (2000), firm's production corresponds to the number of successfully performed problems. The number of problems a firm can solve at each period is limited by its employees' competence and time constraints.

[7] Hereafter we explain that this is an endogenous property of our model.

between the successful employee and the average unemployed should be sufficiently high, a necessary condition for the bonus to be credible, is high heterogeneity in the initial distribution of abilities. Hence, an end of period bonus can be credibly used to motivate agents to perform tasks that reveal their high productivity. Since its use limits to difficult tasks, it can be complemented by a termination contract that guarantees incentives for performing easy tasks. We show that a necessary condition for the simultaneous use of bonus and termination contract is that the expected productivity of insiders is sufficiently larger than the expected productivity of unemployed agents.

If the initial distribution of abilities is not sufficiently heterogeneous, then the rent from keeping an employee is too low. The bonus can not be credibly used and incentives are guaranteed only with an efficiency wage contract. However we show that the set of tasks for which credible incentives are guaranteed is constrained, in the sense that there are tasks for which the firing rule is not credible. Since employees have larger expected productivity than unemployed agents, an insider under-performing a difficult task, could still be more productive than an outsider. The principal prefers to keep the agent rather than to fire him, which destroys the latter's motivation to perform difficult tasks.

To give the intuition of our results we considered above that insiders have larger average ability than unemployed agents. In the model this is endogenously obtained and is due to *ex post* selection by the principal. When a firm owner proposes a termination contract he fires under-performing (less skillful) agents over time. Thus the stationary market equilibrium is characterized by two distributions, inside the firm and on the labor market. They satisfy the property we considered above, namely the average ability of insiders is larger than the one of outsiders. The gap between distributions affects the form of the contract and depends on labor market conditions. For example, a lower exogenous turnover raises the gap of average ability between employed and unemployed agents, thus facilitates the use of the bonus for a larger set of tasks.

The rest of the chapter is organized as follows. Section 2 describes the framework. In Section 3 we present the incentive compatible efficiency wage and bonus contract. In Section 4 we discuss the self-enforceability constraints. In Section 5 we present our main findings. Then we conclude. All proofs are in the Appendix.

2 Framework

Economy We consider an economy composed of a continuum of measure one of firm owners (principals, employers) and a continuum of measure N of potential employees (agents). All parties are endowed with some ability θ. Firm owners are homogeneous and their ability is normalized to 1. Employees are heterogeneous and θ is agent's private information. Let $S(\cdot)$ be the commonly known distribution of workers' ability, $\theta \in [\underline{\theta}, \overline{\theta}]$. We assume that even the most able worker has less ability than a firm owner[8] $(1 > \overline{\theta} > \underline{\theta} \geq 0)$.

An agent is either employed or unemployed at any point in time. The utility of an employee from exerting effort in period t writes as $(W_t - c)$, where W_t is the reward function and c is the effort cost. Any unemployed agent receives an exogenously given unemployment benefit of z. An employer-employee match can be interrupted with probability $(1-\alpha)$ for exogenous reasons. This exogenous turnover rate is the same whatever agent's ability. It causes an employee to enter the unemployment pool.

All parties are infinitely lived and discount the future at a common rate δ.

Production Firm's production corresponds to the number of successfully performed tasks (problems). They are equally valuable to solve and their price is normalized to one. However, some of them are more difficult than others. The

[8] We set aside the endogenous decision to become a firm owner or to remain an employee. However, we consider that only the most able agents become firm owners.

complexity x of a particular task is *ex ante* unknown, but it is drawn from a commonly known distribution. Dealing with a task requires time, succeeding it requires ability and effort. Each agent in the firm has one unit of time by period and spends it completely in dealing with one problem. An employee θ can solve a problem, by expending some effort, only if $x \leq \theta$. If $x > \theta$, he cannot solve it. The effort cost function for an agent θ writes :

$$c(x,\theta) = \begin{cases} c & if \quad x \leq \theta \\ +\infty & if \quad x > \theta \end{cases}$$

Since time is limited, a principal hires employees and delegates them the solution of easy tasks. Thus he spends his time in dealing with problems, that employees are unable to solve. Tasks arrive at the workers' level and it is assumed that any task received by the organization should be solved[9]. Then the number of employees a principal hires and thus the number of problems the organization deals with in each period, is given by the following time constraint :

$$n_t T_t = 1$$

where n_t is the number of employees, and T_t is the proportion of problems, that employees transmit to the principal. T_t is a function of employees' ability and the principal's productivity requirements (*i.e.* the set of problems for which he motivates agents to work), it is determined in Section 5.1.

For the rest of the paper we make the following normalizations. First, problems' complexity is uniformly distributed on the support $[0, 1]$. Hence, θ designs the proportion of problems an agent is able to solve. Second, a principal solves problems

[9] The production function is in the spirit of Garicano (2000). Here production is organized in a two layer knowledge based hierarchy. Each employee receives a problem and solves it if able and supposed to. If not, he transmits it to the principal.

without cost[10].

Output is produced whenever either workers or the firm owner are able to solve a problem, thus the expected output in period t is $y_t = n_t 1$, and a firm owner's expected profit[11] in period t writes as $\Pi_t = n_t(1 - E[W_t])$.

The employer-employee relationship. Some features of a contract are clearly defined, measurable by a third party, thus enforceable. Others are observable inside the firm, but not measurable in a way that could be verified in court. We assume that the difficulty of a given task and whether it has been succeeded or not is not verifiable, thus a legally enforceable contract cannot be written on it. However, it is observed by the employee dealing with the task and the employer. So they can agree on a relational contract, enforced by the possibility of future actions for each of them.

1. At the beginning of each period a firm, with vacant occupations, is matched with agents from the unemployment pool.
2. The principal proposes a contract C to the agents. If the offer is rejected, the agent returns to the unemployment pool and the vacancy remains unoccupied. The contract specifies the reward W_t and implicit[12] performance requirement(s). The reward consists of a fixed wage w_t that the principal pays whatever agent's output and a bonus b_t contingent on employee's performance. The fixed wage is verifiable : a principal cannot renege on its payment. The bonus is paid at the end of a period. Since it is contingent on unverifiable performance, it cannot be enforced by a court.
3. A worker who accepts the contract receives a task, and decides to solve it (if able) or not.

[10] Thus the instantaneous profit of a firm owner that works alone is equal to 1. Indeed, he deals with one task by period, solves it with probability 1 and without effort cost.

[11] Our results are not related to this specification of firm's production function.

[12] Since the performance is not verifiable, the principal cannot commit on explicit (*i.e.* enforceable by a third party) performance requirements.

4. The principal observes the difficulty of each problem, which employee received it and whether it was solved or not. Then he pays or not the bonus. At the end of the period the worker and the firm decide, whether to stay matched for the next period. Exogenous separations occur.

Assumption 3.1. *A principal proposes the same contract to any employee.*

There are two agency problems on the employees' side : moral hazard as effort is not observable, and adverse selection since the agent's type is private information. With Assumption 3.1 we rule out *ex ante* screening and focus on the moral hazard problem. This assumption simplifies the analysis without imposing, in our view, strong restrictions on it. Since contracts contingent on performance are not enforceable, *ex ante* selection is costly for a firm owner. The intuition is that once an agent revealed his type, the principal has the possibility to fire him if he expects a larger profit from an agent of the unemployment pool. Thus to obtain revelation, the menu of contracts proposed by the principal should be such that he gets the same expected profit from any agent whatever his type. These additional constraints increase importantly the cost of screening for the principal[13].

Assumption 3.2. *Labor market is anonymous.*

The reasons why a particular employer-employee match came to end are not observed by any party on the labor market, neither employers, nor employees. We rule out the possibility for firm owners as well as for employees to build external reputation[14].

Assumption 3.3. *Employees are anonymous in the firm.*

[13] However, here pooling is assumed since this informal analysis is not complete. Thus we cannot claim that revelation is obviously sub-optimal. For more complete discussion about the optimal revelation contract under adverse selection without commitment we refer the reader to Laffont and Tirole (1993, pp. 437-460).

[14] This assumption is in the line of MacLeod and Malcomson (1998).

The decision to keep or fire an employee is only based on his performance in the current period. Said differently, a principal does not have the history of employee's previous performance, nor he knows the number of periods the agent spent in the firm. This restriction allows us to focus on the impact of *ex post* selection on the optimal relational contract. However the assumption is not innocuous and we propose a detailed discussion of its implications in Section 6.

Thus the principal proposes the same contract to any employee (Assumption 3.1), and its terms only depend on agent's performance in the current period (Assumption 3.3).

In what follows we begin by separately characterize the incentive compatible termination contract and performance-based bonus. We show that if the principal could write enforceable contracts, he would propose a bonus to motivate agents. Then we present self-enforceability conditions, and discuss the rationale for the simultaneous use of termination contract and bonus.

3 Incentive compatibility

Let $V_t(\theta)$ (*resp.* $V_t^U(\theta)$) represent the expected future utility of an employee (*resp.* unemployed) with type θ. Our analysis focuses on the steady state, thus we drop the t index.

3.1 Termination contract

Under a termination contract the wage paid to an employee is independent of performance. However at the end of the period the employer can terminate the contract if performance is unsatisfactory. Thus a termination contract is specified by a wage-performance pair $(w, \underline{x})$. Where w is the fixed wage and $\underline{x}$ the performance standard to be achieved in order to keep the job[15].

In a setting with heterogeneous agents, there are two potential reasons for an employee to be fired : shirking and lack of performance. Since unsuccessful agents are fired at each period, the choice of performance standard allows *ex post* selection. We first present the incentive compatible fixed wage, then we discuss the selection effect of the associated performance standard.

The incentive compatibility of a termination contract. In each period an employee with ability $\theta < \underline{x}$, is in one of the following situations.

- He receives a problem of difficulty $x \leq \theta$. If he works, he solves it and keeps his job with probability α.
- He receives a problem of difficulty $x > \theta$:
  - if $x < \underline{x}$ the agent is fired for not meeting the standard
  - if $x > \underline{x}$ he keeps his job with probability α.

Then the inter-temporal expected utility of an employee with $\theta < \underline{x}$ who exerts effort is :

$$V(\theta) = w - \theta c + \delta(1-\alpha)V_U(\theta) + \delta\alpha[(1-\underline{x}+\theta)V(\theta) + (\underline{x}-\theta)V_U(\theta)]$$

The expected utility of an agent with $\theta \geq \underline{x}$ who exerts effort writes :

$$V(\theta) = w - \underline{x}c + \delta(1-\alpha)V_U(\theta) + \delta\alpha V(\theta)$$

An unemployed agent receives the unemployment compensation z. At the end of each period, there is an endogenous probability[16] λ to find a new job and quit the unemployment pool. Since labor market is anonymous, this probability is the same

[15] It is not in the principal's interest to fire an agent which performance is satisfactory. Since the expected productivity of an employee is increasing in the difficulty of the transmitted problem, if the principal fires an agent unable to solve a problem $\bar{x} - \epsilon$, then it is in his interest to fire any agent who fails to solve a problem $x < \bar{x} - \epsilon$.

[16] It is determined in Section 5.2.

for any unemployed agent, whatever his employment history. Thus λ is independent of θ. The inter-temporal expected utility of an unemployed agent is :

$$V_U(\theta) = z + \delta[\lambda V(\theta) + (1-\lambda)V_U(\theta)]$$

For a wage to be incentive compatible, the agent's value from working should be larger than his value from shirking. Thus an employee solves a problem only if :

$$w - c + \delta\alpha V(\theta) + \delta(1-\alpha)V_U(\theta) \geq w + \delta V_U(\theta)$$

$$\Leftrightarrow V(\theta) - V_U(\theta) \geq \frac{c}{\alpha\delta} \tag{3.1}$$

Since in our setting the cost of effort is independent of θ, and the probability to be reemployed is the same for any unemployed agent (anonymity), we have the following property.

Lemma 3.1. *The incentive compatible wage does not depend on θ.*

For a formal proof of the lemma, see Appendix 8.1. When the incentive compatibility constraint is binding (*i.e.* $(V(\theta) - V_U(\theta)) = \frac{c}{\alpha\delta}$) and since the reemployment probability is independent of agent's ability level, it is easy to see that the inter-temporal expected utility of an unemployed agent[17] does not depend on θ. Let us now turn to $V(\theta)$. A larger θ (for $\theta < \underline{x}$) corresponds to a higher probability to keep a job (the associated gain is $\delta\alpha(V(\theta) - V_U(\theta))$), but also a higher probability to spend the effort cost c. The marginal gain equals the marginal cost. Hence the inter-temporal expected utility of an employee is also independent of θ. Given these results we simplify the notation as follows, V instead of $V(\theta)$ and V_U instead of $V_U(\theta)$.

[17] $V_U(\theta) = z + \delta V_U(\theta) + \lambda\delta(V(\theta) - V_U(\theta)) \Leftrightarrow V_U(\theta)(1-\delta) = z + \frac{\lambda c}{\alpha}$

Thus the incentive compatible fixed wage[18] is :

$$w = c\underline{x} + \frac{c(1-\delta\alpha)}{\delta\alpha} + (1-\delta)V_U \tag{3.2}$$

Expanding the set of tasks delegated to workers raises the probability that an agent spends effort cost (or the probability to be fired for an agent with type $\theta < \underline{x}$), which reduces V and makes the utility stream when working less attractive for the agent. Thus a larger performance standard should be combined with a higher wage in order to motivate workers.

The selection effect of a performance standard. By choosing a performance standard a principal alters the characteristics of insiders' distribution. Indeed, the choice of $\underline{x}$ sets the proportion of agents who will be fired in each period and replaced by outsiders. Let $F(\theta)$ be the stationary cumulative distribution function of employees' ability. The agents fired in each period are those unable to solve a problem below the performance standard : $v = \int_{\underline{\theta}}^{\underline{x}} F(\theta)d\theta = \int_{\underline{\theta}}^{\underline{x}} (\underline{x}-\theta)f(\theta)d\theta$. Thus for given[19] distribution of outsiders $Q(\theta)$, the steady state distribution of insiders $F(\theta)$ is determined by the following expression.

$$Q(\theta) = \begin{cases} \dfrac{(1-\alpha)F(\theta) + \alpha\displaystyle\int_{\underline{\theta}}^{\theta}(\underline{x}-u)f(u)du}{(1-\alpha+\alpha v)} & \text{if} \quad \theta \in [\underline{\theta};\underline{x}] \\ \dfrac{(1-\alpha)F(\theta)+\alpha v}{(1-\alpha+\alpha v)} & \text{if} \quad \theta \in [\underline{x};\bar{\theta}] \end{cases} \tag{3.3}$$

The steady state distribution of outsiders is composed of employees that quit their job for exogenous reasons, and employees that are fired because their lack of competence has been detected by the principal.

---

[18] This wage motivates any agent to solve any task $x \leq \min\{\theta, \underline{x}\}$.

[19] Each principal is "too small" to have an individual impact on market conditions, thus he takes the distribution of outsiders' ability as given.

Lemma 3.2. *Insiders' distribution of ability dominates the outsiders' distribution, in the sense of first order stochastic dominance.*

3.2 Bonus

A contract is "pure bonus" when the principal motivates agents with a reward based on performance realization. In this case the fixed wage ensures agents' participation (*i.e.* $V(\theta) \geq V_U(\theta)$).

The bonus b is paid to an agent conditionally on performing successfully a task. The contract satisfies the following constraints :

$$\begin{cases} b - c + \alpha\delta V(\theta) + (1-\alpha)\delta V_U(\theta) \geq \delta V_U(\theta) & (IC) \\ V(\theta) \geq V_U(\theta) & (IR) \end{cases}$$

As the principal's profit is decreasing in agents' reward, it is easy to see that he binds both constraints and we have $b = c$, and w such that[20] $V(\theta) = V_U(\theta)$. If such a contract can be enforced, then each employee solves any task he is able to.

Since in the case of termination contract the principal pays rents to the agents, the performance-based pay is the less expensive way to motivate employees. Thus if the principal could write enforceable contracts, he would choose a performance-based pay to provide incentives. However, since performance is not verifiable the employer-employee relationship is steered by a relational contract, and the bonus could only be used to motivate agents to perform a subset of difficult tasks (above some threshold value $\bar{x}$). In the next section we introduce the restrictions imposed on the contract space by the need for the agreement to be self-enforceable.

[20] It is easy to show that w does not depend on θ, and $w = z$. Recall that z is the unemployment benefit.

4 Self-enforceability

4.1 The Contract

The equilibrium self-enforceable contract when there is unemployment is $C^E = (w, b, \underline{x}, \bar{x})$, where w and b are respectively the fixed wage and the bonus, $\underline{x}$ is a minimum performance to be achieved in order to keep a job and $\bar{x}$ defines the lower bound of the set of tasks for which the bonus is used. C^E has the following characteristics.

– For problems $x \geq \bar{x}$ incentives are provided with a performance-based bonus. As we already mentioned it in Section 3, the bonus corresponds to the least costly way for the principal to motivate employees. It is in his interest to adopt this incentive scheme for the largest possible set of tasks. The problem is that the bonus cannot be self-enforced for tasks close to $\underline{\theta}$. Indeed, the principal pays the bonus only if there is a sufficient rent from continuing the employment relationship. The expected productivity of an agent performing $\underline{\theta}$ equals the productivity expected from an outsider. There is no rent from keeping the agent. Then the principal cannot credibly promise the bonus for this task. The bonus is credible only if an agent performs a task above some threshold level $\bar{x} \in]\underline{\theta}, \overline{\theta}]$.

– For problems $x \leq \underline{x}$, incentives are provided with a termination contract. Since the performance-based pay is self-enforceable only for a subset of tasks, the principal can use a termination contract in order to motivate agents to perform easy tasks. An employee that fails to solve a problem $x \leq \underline{x}$ is fired.

– $\bar{x} \geq \underline{x}$

In equilibrium the principal does not duplicate incentive schemes for a given task. To give the intuition for this, let us consider the contrary, $\bar{x} < \underline{x}$. There is a set of problems $\underline{x}$ for which the principal motivates agents with an efficiency wage contract. Moreover the principal pays the bonus to successful agents for the tasks $(\underline{x} - \bar{x})$. This raises the principal's cost without any gain, since the set of solved problems does not change. Thus the firm owner is strictly better off by increasing $\bar{x}$.

– If an agent fails to perform a task $x \geq \bar{x}$, he keeps his job.

Consider a situation in which termination contract and bonus are simultaneously used by the principal, and the incentive compatible efficiency wage is given by (3.2). If the principal fires agents who fail to solve tasks $x \geq \bar{x}$, he destroys the incentives provided with the termination contract to any agent with ability $\theta < \bar{x}$. Indeed, the probability that those employees are fired for lack of performance increases, thus their rent from employment is no more sufficient to guarantee effort provision. A more detailed discussion of this point can be found in Appendix 8.7. Note that this condition is only relevant if the contract proposed by the principal is a mix of bonus and efficiency wage.

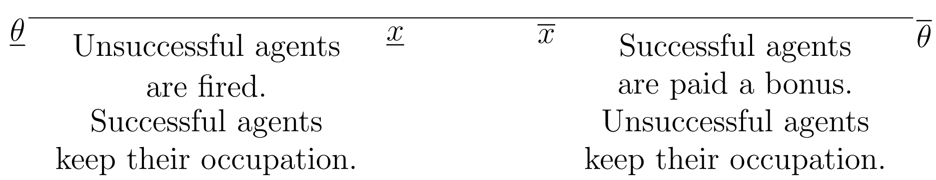


Figure 3.1 : The self-enforceable contract C^E

In the rest of the section we analyze the self-enforceability constraints for the terms of the contract.

4.2 Self-enforceability of the minimum performance standard

The problem of self-enforceability for a termination contract does not concern the fixed wage (a principal can credibly commit on its amount and payment) but the threat to fire an unsuccessful agent. Suppose the performance standard is set too high. Then, as employees are selected, the profit expected from an insider who has not solved a problem below the performance standard could still be larger than the profit expected from an outsider. In such a case a principal cannot credibly commit to fire the unsuccessful employee. Agents anticipate it and do not solve the problem.

The rule to fire any agent performing below the performance standard is self-enforcing, if the profit expected from an insider that fails to solve $\underline{x}$ ($\pi^{in}(\theta \leq \underline{x})$) is lower than the profit expected from an unemployed agent (π^{out}).

$$\pi^{in}(\theta \leq \underline{x}) \leq \pi^{out} \tag{3.4}$$

where

$$\begin{aligned}\pi^{in}(\theta \leq \underline{x}) = \;& \underline{x} - \int_{\underline{\theta}}^{\underline{x}} \frac{(\underline{x}-\theta)f(\theta)}{F(\underline{x})}d\theta - w(\underline{x}) + (1-\alpha)\delta\pi^{out} + \\ & \alpha\delta\Big((1 - \int_{\underline{\theta}}^{\underline{x}} \frac{(\underline{x}-\theta)f(\theta)}{F(\underline{x})}d\theta)\pi^{in}(\theta \leq \underline{x}) + \int_{\underline{\theta}}^{\underline{x}} \frac{(\underline{x}-\theta)f(\theta)}{F(\underline{x})}d\theta\pi^{out}\Big)\end{aligned}$$

$$\pi^{out} = \underline{x} - \int_{\underline{\theta}}^{\underline{x}} (\underline{x}-\theta)q(\theta)d\theta + \int_{\bar{x}}^{\overline{\theta}} (\theta-\bar{x})q(\theta)d\theta(1-b) - w(\underline{x}) + \delta\pi^{out}$$

Since for a given performance standard, the conditional expected productivity of an agent is increasing in the difficulty of the problem that has been transmitted, if the threat of firing is credible for an agent transmitting $\underline{x}$, it will also be credible for an employee who fails to solve $x < \underline{x}$.

After some transformations the self-enforceability condition (3.4) writes :

$$-\big(\underline{x} - \int_{\underline{\theta}}^{\underline{x}} \frac{\theta f(\theta)}{F(\underline{x})}d\theta\big) + Q(\underline{x})\big(\underline{x} - \int_{\underline{\theta}}^{\underline{x}} \frac{\theta q(\theta)}{Q(\underline{x})}d\theta\big) - (1-b)(1-Q(\bar{x}))\big(\frac{\int_{\bar{x}}^{\overline{\theta}} \theta q(\theta)d\theta}{(1-Q(\bar{x}))} - \bar{x}\big) \leq 0$$

In order to simplify the exposition, we adopt the following notation for the rest of the chapter : $l_{in} = \big(\underline{x} - \int_{\underline{\theta}}^{\underline{x}} \frac{\theta f(\theta)}{F(\underline{x})}d\theta\big)$, $l_{out} = \big(\underline{x} - \int_{\underline{\theta}}^{\underline{x}} \frac{\theta q(\theta)}{Q(\underline{x})}d\theta\big)$ and $h = \big(\frac{\int_{\bar{x}}^{\overline{\theta}} \theta q(\theta)d\theta}{(1-Q(\bar{x}))} - \bar{x}\big)$ which is equal[21] to $\big(\frac{\int_{\bar{x}}^{\overline{\theta}} \theta f(\theta)d\theta}{(1-F(\bar{x}))} - \bar{x}\big)$.

[21] $\int_{\bar{x}}^{\overline{\theta}} \frac{\theta f(\theta)}{1-F(\bar{x})}d\theta = \int_{\bar{x}}^{\overline{\theta}} \frac{\theta \frac{q(\theta)(1-\alpha+\alpha v)}{1-\alpha}}{1 - \frac{(1-\alpha+\alpha v)Q(\bar{x})-\alpha v}{1-\alpha}}d\theta = \int_{\bar{x}}^{\overline{\theta}} \frac{\theta q(\theta)}{1-Q(\bar{x})}d\theta$

With the new notation, the constraint (3.4) writes :

$$-\,l_{in} + Q(\underline{x})l_{out} - (1-b)(1-Q(\bar{x}))h \leq 0 \tag{3.5}$$

Note that if there is no difference between insiders' and outsiders' distribution (*i.e.* $F(\theta) = Q(\theta)$ for any θ and $l_{in} = l_{out} = l$), the constraint is $-l(1-Q(\underline{x})) - (1-b)(1-Q(\bar{x}))h \leq 0$, which is always satisfied.

Since insiders are selected and have a larger expected productivity than outsiders, the threat to fire an agent may not be credible for some tasks. It is easy to see that, for example, a performance standard $\underline{x} = \overline{\theta}$ cannot be self-enforced. Indeed the constraint writes $\int_{\underline{\theta}}^{\overline{\theta}} \theta f(\theta)d\theta - \int_{\underline{\theta}}^{\overline{\theta}} \theta q(\theta)d\theta \leq 0$. According to Lemma 3.2 the latter inequality is never satisfied. Therefore the threat to fire the employee is not credible for this task, which destroys agents' incentives to work on it.

4.3 Self-enforceability of the bonus

As a bonus cannot be enforced by a third party the principal should have the right incentives to pay it. There are two possibilities for a principal who reneges on paying the bonus. First, he can fire a successful agent and replace him by an outsider. Second, he can keep the agent without paying the bonus. In the latter case the employee continues to perform tasks for which motivation is guaranteed by the termination contract. Hence, the bonus can only be credible if the ongoing relationship with the employee provides a sufficient rent to the principal and if having employees who perform tasks for which incentives are guaranteed with the bonus is sufficiently valuable for him. We now successively analyze each of those situations.

The principal pays the bonus rather than fire the agent. A firm owner does not pay the bonus unless the expected future gains from employment exceed its value. For given $\bar{x}$, the more difficult the problem solved by an agent, the higher his expected productivity. Thus if the principal pays the bonus when an agent solves a problem $\overline{x}$, he will obviously pay it when the problem solved by an employee is above $\overline{x}$. The self-enforceability constraint states that the expected profit from an insider who solved a problem $\bar{x}$ ($\pi^{in}(\theta \geq \bar{x})$) should be sufficiently high in comparison with the profit expected from an outsider :

$$-b + \alpha\delta\pi^{in}(\theta \geq \bar{x}) \geq \alpha\delta\pi^{out} \tag{3.6}$$

where

$$\pi^{in}(\theta \geq \bar{x}) = \underline{x} + \int_{\bar{x}}^{\overline{\theta}} \frac{(\theta - \bar{x})f(\theta)}{1 - F(\bar{x})} d\theta(1-b) - w(\underline{x}) + (1-\alpha)\delta\pi^{out} + \alpha\delta\pi^{in}(\theta \geq \bar{x})$$

The condition (3.6) becomes :

$$\frac{b(1-\alpha\delta)}{\alpha\delta} - Q(\underline{x})\big(\underline{x} - \frac{\int_{\underline{\theta}}^{\underline{x}} \theta q(\theta) d\theta}{Q(\underline{x})}\big) - Q(\bar{x})\big(\frac{\int_{\bar{x}}^{\overline{\theta}} \theta q(\theta) d\theta}{1 - Q(\bar{x})} - \bar{x}\big)(1-b) \leq 0$$

$$\Leftrightarrow \frac{b(1-\alpha\delta)}{\alpha\delta} - Q(\underline{x})l_{out} - (1-b)Q(\bar{x})h \leq 0 \tag{3.7}$$

This constraint illustrates why it is impossible to use the bonus to provide incentives to perform “easy” tasks. For the bonus to be credible, outsiders should not be close substitutes for an insider who succeeds a problem. If $\bar{x} = \underline{x} = \underline{\theta}$, this condition is never satisfied. For $\underline{x} = \underline{\theta}$ there is no selection, *i.e.* the distribution of insiders is the same as the distribution of outsiders. Thus the expected productivity of an insider who solves a problem close to $\underline{\theta}$ is the same as the expected productivity of an outsider. An agent from the unemployment pool is a perfect substitute for the insider. There is no rent from keeping the successful agent, so the principal cheats on paying the bonus.

The principal pays the bonus rather than renege but keep the agent As mentioned above, another possibility for a principal is to keep a successful agent without paying him the bonus. The following constraint prevents such situations :

$$-b+\alpha\delta\pi^{in}(\theta\geq\bar{x})\geq\alpha\delta\pi^{in}_{dev}(\theta\geq\bar{x}) \tag{3.8}$$

where $\pi^{in}_{dev}(\theta \geq \bar{x})$ is the expected profit from an insider who solved a problem $x=\bar{x}$ after principal's deviation on the payment of the bonus. An employee who has been cheated once on the bonus believes[22] that the principal will continue to cheat. Hence he stops solving problems for which motivation is guaranteed by the bonus. However, the employee still solves problems for which incentives are guaranteed with the termination contract[23].

$$\pi^{in}_{dev}(\theta\geq\bar{x})=\underline{x}-w(\underline{x})+(1-\alpha)\delta\pi^{out}+\alpha\delta\pi^{in}_{dev}(\theta\geq\bar{x})$$

After simplifications (3.8) writes :

$$\frac{b(1-\delta\alpha)}{\alpha\delta}-(1-b)\Big(\frac{\int_{\bar{x}}^{\overline{\theta}}\theta f(\theta)d\theta}{(1-F(\bar{x}))}-\bar{x}\Big)\leq 0$$

$$\Leftrightarrow\frac{b(1-\alpha\delta)}{\alpha\delta}-(1-b)h\leq 0 \tag{3.9}$$

Note that when $\bar{x}$ is close to $\overline{\theta}$, then the expected gain from an employee able to solve these tasks is low, which makes more profitable principal's deviation.

---

[22] There are many patterns of believes consistent with equilibrium. We characterize the set of self-enforcing agreements using the most severe punishment.

[23] It is in the employee's interest to do so, as far as his expected utility when employed and performing tasks below the performance standard $\underline{x}$, is larger than his expected utility if unemployed.

4.4 Self-enforceability of the mix of efficiency wage and bonus

Recall that in an equilibrium with joint use of termination contracts and a bonus the principal does not fire agents who transmit problems $x > \bar{x}$. For this to be credible the profit expected from an insider who is unable of solving a problem $\bar{x}$ ($\pi^{in}(\theta \leq \bar{x})$) must exceed the profit expected from an outsider :

$$\pi^{in}(\theta \leq \bar{x}) \geq \pi^{out} \tag{3.10}$$

where

$$\begin{aligned}\pi^{in}(\theta \leq \bar{x}) = \quad & \underline{x} - \frac{\int_{\underline{\theta}}^{\underline{x}}(\underline{x}-\theta)f(\theta)d\theta}{F(\bar{x})} - w(\underline{x}) + (1-\alpha)\delta\pi^{out} + \\ & \alpha\delta\Big((1 - \frac{\int_{\underline{\theta}}^{\underline{x}}(\underline{x}-\theta)f(\theta)d\theta}{F(\bar{x})})\pi^{in}(\theta \leq \bar{x}) + \frac{\int_{\underline{\theta}}^{\underline{x}}(\underline{x}-\theta)f(\theta)d\theta}{F(\bar{x})}\pi^{out}\Big)\end{aligned}$$

After simplifications the condition (3.10) becomes :

$$\frac{F(\underline{x})}{F(\bar{x})}\big(\underline{x} - \frac{\int_{\underline{\theta}}^{\underline{x}}\theta f(\theta)d\theta}{F(\underline{x})}\big) - Q(\underline{x})\big(\underline{x} - \frac{\int_{\underline{\theta}}^{\underline{x}}\theta q(\theta)d\theta}{Q(\underline{x})}\big) + (1-b)(1-Q(\bar{x}))\big(\frac{\int_{\bar{x}}^{\bar{\theta}}\theta q(\theta)d\theta}{1-Q(\bar{x})} - \bar{x}\big) \leq 0$$

$$\Leftrightarrow \frac{F(\underline{x})}{F(\bar{x})}l_{in} - Q(\underline{x})l_{out} + (1-b)(1-Q(\bar{x}))h \leq 0 \tag{3.11}$$

Without the selection effect on insiders distribution ($l_{in} = l_{out} = l$ and $F(\theta) = Q(\theta)$), the left hand side of the constraint writes as $\Big(lF(\underline{x})(\frac{1}{F(\bar{x})} - 1) + (1-b)(1-Q(\bar{x}))h\Big)$, which is always positive, and the constraint is never satisfied.

Lemma 3.3. *A principal can use simultaneously a termination contract and bonus to motivate an employee to work, only if the expected productivity of the employee is sufficiently larger than the expected productivity of an agent from the unemployment pool.*

5 Equilibrium

5.1 Principal's program

A principal is "too small" to have an impact on market conditions. Thus each firm owner maximizes his profit for a given outside option (V_U) and outsiders' distribution ($Q(\theta)$). Then workers' reservation utility and insiders' and outsiders' distributions are determined consistently in market equilibrium.

The proportion of transmitted problems T under the contract C^E is :
$T = 1 - \underline{x} + \int_{\underline{\theta}}^{\underline{x}} F(\theta)d\theta - \int_{\bar{x}}^{\bar{\theta}} (1 - F(\theta))d\theta.$

A firm owner maximizes his profit :

$$\max_{\underline{x},\bar{x}} \Pi = \frac{1}{T}(1 - w - b\int_{\bar{x}}^{\bar{\theta}} (1 - F(\theta))d\theta)$$

with

$$\begin{cases} b = c \\ w = c\underline{x} + \dfrac{c(1-\alpha\delta)}{\alpha\delta} + (1-\delta)V_U \end{cases}$$

and under the self-enforceability (3.5), (3.7), (3.9), (3.11) constraints.

5.2 The labor market

Employees' outside option. The equilibrium reemployment probability guarantees the equilibrium of flows between the firm and the labor market. The steady state flow into the unemployment pool is $n(1 - \alpha + \alpha v)$. Some agents quit for exogenous reasons $(1 - \alpha)$, others are fired for lack of performance αv. The flow out of the unemployment pool is $\lambda(N - n)$. A proportion λ of the $(N - n)$ unemployed agents finds a job. Thus we obtain : $\lambda = \dfrac{(1-U)}{U}(1-\alpha+\alpha v)$, where $U = \dfrac{N-n}{N}$ corresponds to the unemployment rate.

Abilities distributions. As discussed above, the initial distribution of ability is split into two stationary distributions - inside the firm and on the labor market. In equilibrium the following condition holds :

$$S(\theta) = (1-U)F(\theta) + UQ(\theta) \tag{3.12}$$

The gap between the insiders' and outsiders' productivity plays an important role to characterize the self-enforceable relational contract. As we already explained it this gap is endogenous, and from equation (3.12) we notice that it depends on labor market conditions.

In what follows, we discuss the extend to which agents' heterogeneity affects self-enforceability of termination contracts and performance-based bonus. Furthermore we analyze the extend to which the trade-off between those tools is affected by labor market conditions.

5.3 Pure termination contract

There exist values of the parameters for which a performance-based bonus cannot be self-enforced. Let us denote by m the mean and by σ^2 the variance that characterize the initial distribution of abilities.

Proposition 3.1. *If* $c \geq \min\{\frac{\alpha\delta m}{1-\alpha\delta(1-m)}; \frac{\alpha\delta}{1-\alpha\delta}\Big(\overline{\theta}-m+\frac{\alpha\sigma^2}{1-\alpha+\alpha(\overline{\theta}-m)}\Big)\}$ *then the bonus can not be self-enforced.*

We note that it is impossible to commit on a bonus if the expected surplus from paying it (agent's expected production) is not sufficiently high. More interestingly, the bonus cannot be implemented if the initial distribution of employees is not sufficiently heterogeneous. Indeed a principal is eager to pay the bonus only if the surplus of continuing the relationship with the successful employee is high enough. To say it differently, by solving a difficult task a worker signals that he is more able than the average unemployed agent. This is sufficiently valuable to guarantee the payment of the bonus only if the distribution is not too homogeneous[24].

When Proposition 3.1 holds, the firm's owner motivates employees with a pure termination contract. He maximizes his profit under the self-enforceability constraint of $\underline{x}$:

$$-l_{in} + Q(\underline{x})l_{out} \leq 0$$

In this case the existence of heterogeneity limits the set of self-enforceable efficiency wage contracts, in the sense that a principal cannot credibly commit to fire an insider who fails to solve a difficult task. We show in Appendix 8.5 that there exists some threshold level for $\underline{x}$, denoted by $\hat{\underline{x}}$, such that the principal cannot motivate employees to solve tasks $x > \hat{\underline{x}}$. Since employees are in average more productive than the unemployed, the principal cannot be too demanding with them. Clearly, this restriction on the credible performance requirements is due to the endogenous selection of insiders. The degree of selection depends on the performance standard, but also on labor market conditions, namely the unemployment rate and exogenous turnover.

A higher exogenous turnover raises the incentive cost of the efficiency wage contract. Indeed if agent's probability to quit the firm whatever his performance is larger, it raises the wage the principal should pay to motivate the employee. Thus in a setting the principal is not constrained by the self-enforceability, a higher turnover leads to a decrease of the performance standard. In our setting a higher turnover may play in the opposite sense.

Proposition 3.2. *A higher turnover (i.e. lower α) softens the self-enforceability constraint and thus increases $\hat{\underline{x}}$.*

[24] To illustrate our purpose assume that initially ability is uniformly distributed on the support $[m-\epsilon, m+\epsilon]$. Then the condition $\frac{\alpha\delta}{1-\alpha\delta}\left(\overline{\theta} - m + \frac{\alpha\sigma^2}{1-\alpha+\alpha(\overline{\theta}-m)}\right)$ writes as $\frac{\epsilon(1-\alpha)+(4/3)\alpha\epsilon^2}{1-\alpha+\alpha\epsilon}$. The latter only depends on ϵ and increases with its value.

A higher turnover reduces the degree of heterogeneity between the distribution of insiders and that of outsiders. First, when the proportion of agents who quit the firm for exogenous reasons is larger, insiders are less selected. The intuition is that among the employees who successfully performed the required tasks, larger proportion quits the firm and is replaced by unemployed agents. Second, in the unemployment pool, the proportion of unemployed for exogenous reasons (rather than for lack of performance) increases, which improves outsiders' distribution. A lower gap in the expected productivity of insiders and outsiders expands the set of tasks for which the threat to fire an unsuccessful insider is credible.

5.4 Efficiency wage and bonus

In this section we discuss the characteristics of the incentive mix, when the principal uses both termination contract and performance-based bonus. Performing a difficult task signals high ability, and if outsiders are not close substitutes to insiders, the bonus payment may be credible. However the principal can use it to motivate effort only for a fraction of tasks, the more difficult ones. Thus it may be complemented by a termination contract, which motivates agents to perform easy tasks.

It is in the principal's interest to use the bonus as a motivation tool for the largest possible set of tasks ; he should choose the lower $\bar{x}$, compatible with the self-enforceability constraints. Hereafter we analyze which one of the following constraints is the most stringent, and thus determines the principal's decision.

$$\begin{cases} -l_{in} + Q(\underline{x})l_{out} - (1-c)(1-Q(\bar{x}))h \leq 0 & (3.5) \\ \dfrac{c(1-\alpha\delta)}{\alpha\delta} - (1-c)h - Q(\underline{x})l_{out} + (1-c)(1-Q(\bar{x}))h \leq 0 & (3.7) \\ \dfrac{c(1-\alpha\delta)}{\alpha\delta} - (1-c)h \leq 0 & (3.9) \\ \dfrac{F(\underline{x})}{F(\bar{x})} l_{in} - Q(\underline{x})l_{out} + (1-c)(1-Q(\bar{x}))h \leq 0 & (3.11) \end{cases}$$

Constraint (3.11) guarantees that the expected productivity of an outsider does not exceed the expected productivity of an insider who is unable to solve $\bar{x}$. If this constraint is satisfied and if the surplus of tasks under bonus contract is sufficiently high to justify its payment (*i.e.* constraint (3.9) is satisfied), then it is never in the principal's interest to fire a successful insider and replace him by an outsider. To say it differently, if the principal uses a mix of efficiency wage and performance-based bonus, constraint (3.7) never binds.

The constraint (3.9) rewrites (by using (3.12)) as :

$$\frac{c(1-\alpha\delta)}{\alpha\delta} - (1-c)\Big(\int_{\overline{x}}^{\overline{\theta}} \frac{\theta s(\theta)}{1-S(\overline{x})} d\theta - \overline{x}\Big) \leq 0$$

For a large set of initial ability distributions, the left hand side of the constraint is increasing[25] in $\bar{x}$ (*i.e.* larger $\bar{x}$ makes the constraint more difficult to satisfy). When the bonus is used for a small set of tasks, then the expected gain from an agent able to solve these tasks is smaller, which makes more profitable principal's deviation.

Conversely, the left hand side of (3.11) decreases with $\bar{x}$. So if there exists a solution for $\bar{x}$ that satisfies both constraints, the principal chooses the lowest possible $\bar{x}$, constraint (3.11) obviously binds.

Finally, if the expected productivity of an outsider equals the expected productivity of an insider who is unable to solve $\bar{x}$, it is obviously larger than the expected productivity of an insider who cannot solve $\underline{x}$ (since $\bar{x} \geq \underline{x}$). So if (3.11) binds, (3.5) is satisfied.

The principal maximizes his profit, with respect to $\underline{x}$, and $\bar{x}$ is given by the constraint :

$$\frac{F(\underline{x})}{F(\bar{x})} l_{in} - Q(\underline{x}) l_{out} + (1-c)(1-Q(\bar{x}))h = 0$$

---

[25] As shown by Bagnoli and Bergstrom (2005), a sufficient condition for $\Big(\frac{\int_{\bar{x}}^{\overline{\theta}} \theta s(\theta) d\theta}{(1-S(\bar{x}))} - \bar{x}\Big)$ to be monotone decreasing in $\bar{x}$ is that the density function $s(\theta)$ and the reliability function $\int_{\bar{x}}^{\overline{\theta}}(1-S(\theta))d\theta$ are log-concave. For a list of the distributions that satisfy these conditions, we refer the reader to Bagnoli and Bergstrom (2005).

When the constraint is satisfied with equality, an increase of $\underline{x}$ raises $\bar{x}$. Thus in equilibrium both incentive tools are substitutes. A more extensive use of the bonus reduces the set of tasks for which the principal motivates the agent through the efficiency wage.

Proposition 3.3. *A lower turnover (i.e. high α) expands the set of tasks for which the bonus is credible (i.e. $\bar{x}$ decreases).*

In markets with low turnover, the bonus can be more extensively used by the principal. There are two reasons for that. First, a lower turnover increases the probability of continuation of the relationship, thus increasing the cost of reneging on the agreement. Second, it expands the gap between insiders' and outsiders' productivity, which is the effect highlighted in Proposition 3.3. As already discussed in Section 5.3, a lower turnover improves insiders' distribution and worsens outsiders' distribution, thus increasing the productivity gap between the average employee and unemployed agent, which in turn relaxes constraint (3.11) and reduces $\bar{x}$. In fine the positive effect of lower turnover on the utilization of the bonus is twofold. It is used for a larger set of tasks, furthermore insiders are better selected, thus problems for which the bonus is paid are solved more often.

6 Discussion

This chapter provides some new elements in the understanding of incentive schemes used by managers. We derive conditions for the simultaneous use of termination contracts and bonus payments in order to motivate employees to work. We show that this is possible only if there is a difference in the expected productivity of an employee and an unemployed agent. The gap in expected productivities of employees and unemployed, in our setting, can be driven by : either the impact of *ex post* selection on insiders' and outsiders' distributions or agent's seniority. In the chapter we focus on the first effect and set the second one aside (with

Assumption 3.3). Over time unsuccessful employees' are fired and replaced, insiders' and outsiders' stationary distributions are endogenously obtained, which generates the difference in the expected productivity of employed and unemployed agents.

Let us now briefly discuss the potential effect of seniority. To expose our intuitions we assume that the pool of unemployed has the same characteristics over time (set aside the endogenous selection). An agent who has been in the firm for a long time has larger expected productivity than an unemployed one. The reason for that is related on the fact that an agent with longer career in the firm is one that succeeded successive tests, *i.e.* an employee that has not been fired for lack of performance. When an agent arrives in the firm the bonus cannot be credibly proposed since his expected productivity is the same as that of an unemployed. However with employee's seniority bonus could be used to motivate the agent to perform some difficult tasks. The larger the number of periods the agent spends in the firm the larger the set of tasks for which the bonus can be credibly used. We hope that a more precise characterization of the optimal contract with seniority would be the subject of a future work. However we think that in a model that would take into account both the endogenous evolution of distributions of insiders and outsiders, and the effect of seniority, the findings of this chapter would be still present.

7 References

Bagnoli M. and T. Bergstrom (2005), “Log Concave Probability and Its Applications”, *Economic Theory*, vol. 26, pp. 445-469.

Baker G., R. Gibbons and K.J. Murphy (1994), “Subjective performance measure in Optimal Incentive Contracts”, *Quarterly Journal of Economics*, vol. 109, pp. 1125-1156.

Baker G., M. Gibbs and B. Holmström (1994), “The Wage Policy of a Firm”, *Quarterly Journal of Economics*, vol. 109, pp. 921-955.

Bull C. (1987), “The Existence of Self-Enforcing Implicit Contracts”, *Quarterly Journal of Economics*, vol. 102, pp. 147-159.

Cappelli P. and K. Chauvin (1991), “An Interplant Test of the Efficiency Wage Hypothesis”, *The Quarterly Journal of Economics*, vol. 106, pp. 769-787.

Garicano L. (2000), “Hierarchies and the Organization of Knowledge in Production”, *The Journal of Political Economy*, vol. 108, pp. 874-904.

Hayes, R. and S. Schaefer, (2000), “Implicit Contracts and the Explanatory Power of Top Executive Compensation for Future Performance,” *RAND Journal of Economics*, Vol. 31, pp. 273-293.

Laffont J.J. and J. Tirole (1993), “A Theory of Incentives in Procurement and Regulation”, *MIT press*

Levin J. (2003), “Relational Incentive Contracts”, *American Economic Review*, vol. 93, pp. 835-847.

MacLeod W.B. and J.M. Malcomson (1989), “Implicit Contracts, Incentive Compatibility, and Involuntary Unemployment”, *Econometrica*, vol. 57, pp. 447-480.

MacLeod W.B. and J.M. Malcomson (1998), “Motivation and Markets”, *American Economic Review*, vol. 88, pp. 388-411.

MacLeod W.B. and D. Parent (1997), "Jobs Characteristics and the Form of Compensation", Mimeo, University of Southern California.

Shapiro C. and J. Stiglitz (1984), "Equilibrium Unemployment as a Worker Discipline Device", *American Economic Review*, vol. 74, pp. 433-444.

8 Appendix

8.1 Proof of Lemma 3.1

The incentive compatibility constraint for the fixed wage of a termination contract writes :

$$V(\theta) - V_U(\theta) \geq \frac{c}{\alpha\delta}$$

For an agent with type $\theta < \underline{x}$. The difference between the expected utility when employed and the expected utility when unemployed is $V(\theta) - V_U(\theta) = \frac{w - z - \theta c}{1 - \alpha\delta(1 - \underline{x} + \theta) + \lambda\delta}$. Thus the incentive compatibility constraint corresponds to :

$$\frac{w - z - \theta c}{1 - \alpha\delta(1 - \underline{x} + \theta) + \lambda\delta} \geq \frac{c}{\alpha\delta} \Leftrightarrow w \geq \frac{c(1 - \alpha\delta)}{\alpha\delta} + c\underline{x} + z + \frac{c\lambda}{\alpha}$$

For an agent with type $\theta \geq \underline{x}$. Now we have :

$$V(\theta) - V_U(\theta) = \frac{w - z - \underline{x}c}{1 - \alpha\delta + \lambda\delta}$$

$$\frac{w - z - \underline{x}c}{1 - \alpha\delta + \lambda\delta} \geq \frac{c}{\alpha\delta} \Leftrightarrow w \geq \frac{c(1 - \alpha\delta)}{\alpha\delta} + c\underline{x} + z + \frac{c\lambda}{\alpha}$$

Thus the incentive compatible wage does not depend on θ.

8.2 Stationary distributions

The distribution of insiders at the beginning of a period is $F(\theta)$. $\tilde{F}(\theta)$ is the modified distribution after departures (for exogenous and endogenous reasons) took place. Vacancies are fulfilled to agents from the unemployment pool, with cumulative distribution function $Q(\theta)$.

Stationarity condition writes :

$$\underbrace{(1-\alpha+\alpha v)}_{\text{proportion of hired agents}} Q(u) + \underbrace{\alpha(1-v)}_{\text{proportion of remaining insiders}} \tilde{F}(u) = F(u) \tag{3.13}$$

Where $\tilde{F}(u)$ is :

$$\tilde{F}(u) = \begin{cases} \dfrac{\displaystyle\int_{\underline{\theta}}^{u} (1-(\underline{x}-\theta))f(\theta)d\theta}{1-v} & if \quad u \in [\underline{\theta};\underline{x}] \\ \dfrac{F(u)-v}{1-v} & if \quad u \in [\underline{x};\bar{\theta}] \end{cases} \tag{3.14}$$

Thus from (3.13) and (3.14) we obtain :

$$Q(u) = \begin{cases} \dfrac{F(u)(1-\alpha)+\alpha\displaystyle\int_{\underline{\theta}}^{u} (\underline{x}-\theta)f(\theta)d\theta}{1-\alpha+\alpha v} & if \quad u \in [\underline{\theta};\underline{x}] \\ \dfrac{F(u)(1-\alpha)+\alpha v}{1-\alpha+\alpha v} & if \quad u \in [\underline{x};\bar{\theta}] \end{cases}$$

8.3 Proof of Lemma 3.2

The case of $\theta \geq \underline{x}$:

$$Q(\theta) - F(\theta) = \frac{(1-\alpha)F(\theta)+\alpha v}{1-\alpha+\alpha v} - F(\theta) = \frac{\alpha v(1-F(\theta))}{1-\alpha+\alpha v}$$

It is easy to see that

- $\dfrac{\alpha v(1-F(\theta))}{1-\alpha+\alpha v} > 0$, for any $\theta < \overline{\theta}$
- $\dfrac{\alpha v(1-F(\theta))}{1-\alpha+\alpha v} = 0$ for $\theta = \overline{\theta}$

The case of $\theta < \underline{x}$:

$$Q(\theta) - F(\theta) = \frac{(1-\alpha)F(\theta)+\alpha\int_{\underline{\theta}}^{\theta}(\underline{x}-u)f(u)du}{1-\alpha+\alpha v} - F(\theta) = \frac{\alpha\int_{\underline{\theta}}^{\theta}(\underline{x}-u-v)f(u)du}{1-\alpha+\alpha v}$$

$\int_{\underline{\theta}}^{\theta}(\underline{x}-u-v)f(u)du$ increases with θ when $\theta < \underline{x}-v$ and decreases with θ when $\theta > \underline{x}-v$. So if $\int_{\underline{\theta}}^{\theta}(\underline{x}-u-v)f(u)du \geq 0$ for $\theta = 0$ and $\theta = \underline{x}$, it is also positive for any $\theta \in]\underline{\theta}, \underline{x}[$.

- For $\theta = x_w$, $\int_{\underline{\theta}}^{x_w}(x_w-u-v_w)f(u)du = (1-F(x_w))v_w > 0$.
- For $\theta = \underline{\theta}$, $\int_{\underline{\theta}}^{\underline{\theta}}(x_w-u-v_w)f(u)du = 0$.

8.4 Proof of Proposition 3.1

The condition $c \geq \dfrac{\alpha\delta m}{1-\alpha\delta(1-m)}$

It is obtained from constraint (3.9) :

$$\frac{c(1-\alpha\delta)}{(1-c)\alpha\delta} - \int_{\bar{x}}^{\overline{\theta}} \frac{(\theta-\bar{x})s(\theta)}{1-S(\bar{x})} d\theta \leq 0$$

We assume that the initial distribution is such that the left hand side increases with $\bar{x}$. A sufficient condition for this is that the density function $s(\theta)$ and the reliability function $\int_{\bar{x}}^{\overline{\theta}}(1-S(\theta))d\theta$ are log-concave. See Bagnoli and Bergstrom (2005) for a list of distributions that satisfy those conditions. Thus if the constraint is not satisfied for $\bar{x} = 0$ it will never be satisfied. This we obtain the following condition :

$\dfrac{c(1-\alpha\delta)}{(1-c)\alpha\delta} - \displaystyle\int_{\underline{\theta}}^{\overline{\theta}} \theta s(\theta)d\theta > 0 \Leftrightarrow c \geq \dfrac{\alpha\delta m}{1-\alpha\delta(1-m)}$

The condition $c \geq \dfrac{\alpha\delta}{1-\alpha\delta}\Big(\overline{\theta} - m + \dfrac{\alpha\sigma^2}{1-\alpha+\alpha(\overline{\theta}-m)}\Big)$

Constraint (3.7) is :

$$\frac{c(1-\alpha\delta)}{\alpha\delta} - Q(\underline{x})\big(\underline{x} - \frac{\int_{\underline{\theta}}^{\underline{x}}\theta q(\theta)d\theta}{Q(\underline{x})}\big) - (1-c)Q(\bar{x})\big(\frac{\int_{\bar{x}}^{\overline{\theta}}\theta q(\theta)d\theta}{1-Q(\bar{x})} - \bar{x}\big) \leq 0$$

We are interested on the case, this constraint is not satisfied even in the most favorable situation.

First, the left hand side is decreasing in $\underline{x}$. Thus for $\underline{x} = \bar{x}$ we have :

$$\frac{c(1-\alpha\delta)}{\alpha\delta} - (1-c)\frac{Q(\bar{x})}{1-Q(\bar{x})}\int_{\bar{x}}^{\overline{\theta}}(1-Q(\theta))d\theta - \int_{\underline{\theta}}^{\bar{x}}(\bar{x}-\theta)q(\theta)d\theta \leq 0$$

The latter is the most easy to satisfy for $\bar{x} = \overline{\theta}$:

$$\frac{c(1-\alpha\delta)}{\alpha\delta} - \int_{\underline{\theta}}^{\overline{\theta}}(\overline{\theta}-\theta)q(\theta)d\theta \leq 0 \Leftrightarrow \frac{c(1-\alpha\delta)}{\alpha\delta} - \int_{\underline{\theta}}^{\overline{\theta}}\frac{(\overline{\theta}-\theta)s(\theta)(1-\alpha+\alpha(\overline{\theta}-\theta))}{1-\alpha+\alpha v+U\alpha(\overline{\theta}-\theta-v)}d\theta \leq 0$$

Finally we can show that the constraint is the easiest to satisfy for $U \to 0$. Thus the constraint becomes :

$$\frac{c(1-\alpha\delta)}{\alpha\delta} - \int_{\underline{\theta}}^{\overline{\theta}}\frac{(\overline{\theta}-\theta)s(\theta)(1-\alpha+\alpha(\overline{\theta}-\theta))}{1-\alpha+\alpha v}d\theta \leq 0$$

where $v = \int_{\underline{\theta}}^{\overline{\theta}}(\overline{\theta}-\theta)s(\theta)d\theta = \overline{\theta}-m$ and $\int_{\underline{\theta}}^{\overline{\theta}}(\overline{\theta}-\theta)^2 s(\theta)d\theta = (\sigma^2+(\overline{\theta}-m)^2)$. We replace in $\frac{c(1-\alpha\delta)}{\alpha\delta} - \int_{\underline{\theta}}^{\overline{\theta}}\frac{(\overline{\theta}-\theta)s(\theta)(1-\alpha+\alpha(\overline{\theta}-\theta))}{1-\alpha+\alpha v}d\theta \leq 0$, which after transformation equals $c \leq \frac{\alpha\delta}{1-\alpha\delta}\left(\overline{\theta}-m+\frac{\alpha\sigma^2}{1-\alpha+\alpha(\overline{\theta}-m)}\right)$.

So if $c \geq \frac{\alpha\delta}{1-\alpha\delta}\left(\overline{\theta}-m+\frac{\alpha\sigma^2}{1-\alpha+\alpha(\overline{\theta}-m)}\right)$, the constraint (3.7) is never satisfied.

8.5 Proof that some performance standards are not self-enforceable

We show that the self-enforceability constraint for $\underline{x}$ is satisfied for some values and that there exists some $\underline{\hat{x}}$ above which performance standards are not credible. We proceed in two steps :

- First we show that the constraint is binding for $\underline{x} = \underline{\theta} \equiv 0$ and not satisfied for $\underline{x} = \overline{\theta}$.
- Second we show that it is decreasing in $\underline{x}$, for $\underline{x} \to 0$.

First step : The constraint is :

$$B \equiv -\big(\underline{x} - \int_0^{\underline{x}} \frac{\theta f(\theta)}{F(\underline{x})} d\theta\big) + Q(\underline{x})\big(\underline{x} - \int_0^{\underline{x}} \frac{\theta q(\theta)}{Q(\underline{x})} d\theta\big) \leq 0$$

– For $\underline{x} = \overline{\theta}$, $B = \displaystyle\int_0^{\overline{\theta}} \theta f(\theta) d\theta - \int_0^{\overline{\theta}} \theta q(\theta) d\theta$. Under Lemma 3.2 we have $B > 0$, the constraint is not satisfied.

– For $\underline{x} = 0$, $B = 0$

Second step : We now show that B is decreasing in $\underline{x}$, for $\underline{x} \to 0$.

Let us first assume that in the firm and on the labor market there is the same distribution $F(\theta)$ (there is no specification on the distribution function $F(\cdot)$). In this case the constraint writes :

$$\tilde{B} \equiv -\big(\underline{x} - \int_0^{\underline{x}} \frac{\theta f(\theta)}{F(\underline{x})} d\theta\big) + F(\underline{x})\big(\underline{x} - \int_0^{\underline{x}} \frac{\theta f(\theta)}{F(\underline{x})} d\theta\big) \leq 0$$

It is satisfied for any $\underline{x}$. For $\underline{x} = 0$, $\tilde{B} = 0$. So the derivative of $\tilde{B}$ with respect to $\underline{x}$, for $\underline{x} = 0$ should be negative. We have $\displaystyle\frac{\partial \tilde{B}}{\partial \underline{x}} = -(1 - F(\underline{x})) + \frac{f(\underline{x})}{F(\underline{x})}(\underline{x} - \int_0^{\underline{x}} \frac{\theta f(\theta)}{F(\underline{x})})\Big|_{\underline{x}=0} < 0$

Let us now turn to the case of different distributions, one on the market, and one inside the firm. After simplifications the derivative of the constraint writes :

$$\frac{\partial B}{\partial \underline{x}} = -\frac{Q(\underline{x})}{F(\underline{x})}\Big(\frac{\frac{\partial v}{\partial \underline{x}}}{F(\underline{x})} - F(\underline{x})\Big) + \frac{f(\underline{x})}{F(\underline{x})}\Big(\underline{x} - \int_0^{\underline{x}} \frac{\theta f(\theta)}{F(\underline{x})}\Big)$$

$\displaystyle\frac{\partial v}{\partial \underline{x}} = \int_0^{\underline{x}} \frac{f(\theta)(1 - \alpha + \alpha v)}{(1 - \alpha + \alpha(\underline{x} - \theta))} d\theta \geq \frac{F(\underline{x})(1 - \alpha + \alpha v)}{(1 - \alpha + \alpha \underline{x})}$ and from Lemma 3.2 we know that $Q(\underline{x}) \geq F(\underline{x})$. It follows that :

$$\frac{\partial B}{\partial \underline{x}} \leq -\Big(\frac{1 - \alpha + \alpha v}{1 - \alpha + \alpha \underline{x}} - F(\underline{x})\Big) + \frac{f(\underline{x})}{F(\underline{x})}\Big(\underline{x} - \int_0^{\underline{x}} \frac{\theta f(\theta)}{F(\underline{x})}\Big)$$

For $\underline{x} \to 0$ the latter expression converges to $\displaystyle\frac{\partial \tilde{B}}{\partial \underline{x}}\Big|_{\underline{x}=0}$ so we should have $\displaystyle\frac{\partial B}{\partial \underline{x}}\Big|_{\underline{x}=0} < 0$.

8.6 Proof of Propositions 3.2 and 3.3

Before proceeding to the proof of the propositions, we derive the expressions of $f(\theta)$ and $q(\theta)$ as functions of the initial distribution of abilities $s(\theta)$:

$$f(\theta) = \begin{cases} \dfrac{(1-\alpha+\alpha v)s(\theta)}{(1-\alpha+\alpha v)+U\alpha(\underline{x}-\theta-v)} & \text{if} \quad \theta \in [\underline{\theta}; \underline{x}] \\ \dfrac{(1-\alpha+\alpha v)s(\theta)}{(1-\alpha+\alpha v)-U\alpha v} & \text{if} \quad \theta \in [\underline{x}; \bar{\theta}] \end{cases}$$

$$q(\theta) = \begin{cases} \dfrac{(1-\alpha+\alpha(\underline{x}-\theta))s(\theta)}{(1-\alpha+\alpha v)+U\alpha(\underline{x}-\theta-v)} & \text{if} \quad \theta \in [\underline{\theta}; \underline{x}] \\ \dfrac{(1-\alpha)s(\theta)}{(1-\alpha+\alpha v)-U\alpha v} & \text{if} \quad \theta \in [\underline{x}; \bar{\theta}] \end{cases}$$

Equations (3.3) and (3.12) are used.

Recall that $v = \displaystyle\int_{\underline{\theta}}^{\underline{x}} F(\theta)d\theta$. It is immediate to see that $v < \underline{x}F(\underline{x}) < \underline{x}$.

Hereafter, we consider $\underline{\theta} = 0$.

8.6.1 Proof of Proposition 3.2

The self-enforceability constraint for termination contract is :

$$-\frac{\int_0^{\underline{x}}(\underline{x}-\theta)f(\theta)d\theta}{F(\underline{x})} + \int_0^{\underline{x}}(\underline{x}-\theta)q(\theta)d\theta \leq 0$$

The derivative with respect to α writes :

$$-\frac{1}{F(\underline{x})}\int_0^{\underline{x}}(\underline{x}-\theta)\frac{df(\theta)}{d\alpha}d\theta + \frac{v}{F(\underline{x})^2}\int_0^{\underline{x}}\frac{df(\theta)}{d\alpha}d\theta + \int_0^{\underline{x}}(\underline{x}-\theta)\frac{dq(\theta)}{d\alpha}d\theta$$

We adopt the following notation :

$$A \equiv -\frac{1}{F(\underline{x})}\int_0^{\underline{x}}(\underline{x}-\theta)\frac{df(\theta)}{d\alpha}d\theta + \frac{v}{F(\underline{x})^2}\int_0^{\underline{x}}\frac{df(\theta)}{d\alpha}d\theta$$

$$B \equiv \int_0^{\underline{x}}(\underline{x}-\theta)\frac{dq(\theta)}{d\alpha}d\theta$$

In what follows we show that each of these terms is positive.

Proof that $B > 0$

$$B = \int_0^{\underline{x}} \frac{(1-U)s(\theta)(\underline{x}-\theta-v)(\underline{x}-\theta)}{((1-\alpha+\alpha v)+U\alpha(\underline{x}-\theta-v))^2}d\theta - \frac{\partial v}{\partial \alpha}\int_0^{\underline{x}} \frac{(1-U)s(\theta)\alpha(1-\alpha+\alpha(\underline{x}-\theta))(\underline{x}-\theta)}{((1-\alpha+\alpha v)+U\alpha(\underline{x}-\theta-v))^2}d\theta$$

– The first term rewrites as follows : $\dfrac{(1-U)}{1-\alpha+\alpha v}\displaystyle\int_0^{\underline{x}} \frac{f(\theta)(\underline{x}-\theta-v)(\underline{x}-\theta)}{((1-\alpha+\alpha v)+U\alpha(\underline{x}-\theta-v))}d\theta$
We notice that for low (resp high) θ, $f(\theta)(\underline{x}-\theta-v)$ is positive (resp negative). Since $\dfrac{(\underline{x}-\theta)}{((1-\alpha+\alpha v)+U\alpha(\underline{x}-\theta-v))}$ is decreasing in θ the positive terms are highly weighted. Thus if the expression is positive when all the terms are equally weighted, it will also be positive in the case of highly weighted positive terms. With equally weighted terms we have :

$$\begin{aligned}\frac{(1-U)k}{1-\alpha+\alpha v}\int_0^{\underline{x}} f(\theta)(\underline{x}-\theta-v)d\theta &= \frac{(1-U)k}{1-\alpha+\alpha v}v(1-F(\underline{x})) \\ &> 0\end{aligned}$$

where k is a constant. So,

$$\frac{(1-U)}{1-\alpha+\alpha v}\int_0^{\underline{x}} \frac{f(\theta)(\underline{x}-\theta-v)(\underline{x}-\theta)}{((1-\alpha+\alpha v)+U\alpha(\underline{x}-\theta-v))}d\theta > 0$$

– Now we show that $\dfrac{\partial v}{\partial \alpha} < 0$.

$$\begin{aligned}\frac{\partial v}{\partial \alpha}\gamma &= -\int_0^{\underline{x}} \frac{U(\underline{x}-\theta)s(\theta)(\underline{x}-\theta-v)}{((1-\alpha+\alpha v)+U\alpha(\underline{x}-\theta-v))^2}d\theta \\ &< 0\end{aligned}$$

where $\gamma = \left(\dfrac{(1-\alpha)}{1-\alpha+\alpha v} + \displaystyle\int_0^{\underline{x}} \frac{(\underline{x}-\theta)s(\theta)\alpha(1-U)(1-\alpha+\alpha v)}{((1-\alpha+\alpha v)+U\alpha(\underline{x}-\theta-v))^2}d\theta\right)$, which is positive.

To show that $-\displaystyle\int_0^{\underline{x}} \frac{U(\underline{x}-\theta)s(\theta)(\underline{x}-\theta-v)}{((1-\alpha+\alpha v)+U\alpha(\underline{x}-\theta-v))^2}d\theta < 0$ we apply the same kind of argument as above.

Thus we have $B > 0$.

Proof that $A > 0$

$$A = \underbrace{-\frac{1}{F(\underline{x})^2}\int_0^{\underline{x}} \frac{\partial f(\theta)}{\partial \alpha}((\underline{x}-\theta)F(\underline{x})-v)d\theta}_{\equiv a} \underbrace{-\frac{1}{F(\underline{x})^2}\frac{\partial v}{\partial \alpha}\int_0^{\underline{x}} \frac{\partial f(\theta)}{\partial v}((\underline{x}-\theta)F(\underline{x})-v)d\theta}_{\equiv b}$$

– Proof of $a > 0$

$$\begin{aligned} a &= \frac{1}{F(\underline{x})^2}\int_0^{\underline{x}} \frac{Us(\theta)(\underline{x}-\theta-v)((\underline{x}-\theta)F(\underline{x})-v)}{((1-\alpha+\alpha v)+U\alpha(\underline{x}-\theta-v))^2}d\theta \\ &= \frac{1}{F(\underline{x})^2(1-\alpha+\alpha v)}\int_0^{\underline{x}} \frac{Uf(\theta)(\underline{x}-\theta-v)((\underline{x}-\theta)F(\underline{x})-v)}{((1-\alpha+\alpha v)+U\alpha(\underline{x}-\theta-v))}d\theta \\ &\geq \frac{U}{F(\underline{x})^2(1-\alpha+\alpha v)}\int_0^{\underline{x}-v} \frac{f(\theta)(\underline{x}-\theta-v)((\underline{x}-\theta)F(\underline{x})-v)}{((1-\alpha+\alpha v)+U\alpha(\underline{x}-\theta-v))}d\theta \end{aligned}$$

In the last expression the positive terms are highly weighted. So it is sufficient to show that the expression is non negative when all terms are equally weighted :

$$\begin{aligned} \int_0^{\underline{x}-v} f(\theta)(F(\underline{x})(\underline{x}-\theta)-v)d\theta &= vF(\underline{x})F(\underline{x}-v) + F(\underline{x})(v - \int_{\underline{x}-v}^{\underline{x}} F(\theta)d\theta) - vF(\underline{x}-v) \\ &\geq vF(\underline{x})F(\underline{x}-v) + F(\underline{x})(v - F(\underline{x})v) - vF(\underline{x}-v) \\ &\geq v(1-F(\underline{x}))(F(\underline{x})-F(\underline{x}-v)) \\ &> 0 \end{aligned}$$

This implies that $a > 0$.

– Proof of $b > 0$

$$b = -\frac{U\alpha}{F(\underline{x})^2(1-\alpha+\alpha v)}\frac{\partial v}{\partial \alpha}\int_0^{\underline{x}} \frac{f(\theta)(1-\alpha+\alpha(\underline{x}-\theta))(F(\underline{x})(\underline{x}-\theta)-v)}{(1-\alpha+\alpha v)+U\alpha(\underline{x}-\theta-v)}d\theta$$

As we have already shown $\frac{\partial v}{\partial \alpha} < 0$, so the sign of b is given by the expression $\int_0^{\underline{x}} \frac{f(\theta)(1-\alpha+\alpha(\underline{x}-\theta))(F(\underline{x})(\underline{x}-\theta)-v)}{(1-\alpha+\alpha v)+U\alpha(\underline{x}-\theta-v)}d\theta \geq k\int_0^{\underline{x}} f(\theta)(F(\underline{x})(\underline{x}-\theta)-v)d\theta = 0$
$\Rightarrow b > 0$

8.6.2 Proof of Proposition 3.3

The constraint writes :

$$\frac{\int_0^{\underline{x}}(\underline{x}-\theta)f(\theta)d\theta}{F(\bar{x})} - \int_0^{\underline{x}}(\underline{x}-\theta)q(\theta)d\theta + (1-b)\int_{\bar{x}}^{\bar{\theta}}(1-Q(\theta))d\theta \leq 0$$

The derivative with respect to α is :

$$\underbrace{\frac{1}{F(\bar{x})^2}\Big(\int_0^{\underline{x}}(\underline{x}-\theta)F(\bar{x})\frac{df(\theta)}{d\alpha}d\theta - v\int_0^{\bar{x}}\frac{df(\theta)}{d\alpha}d\theta\Big)}_{\equiv J}\underbrace{-\int_0^{\underline{x}}(\underline{x}-\theta)\frac{dq(\theta)}{d\alpha}d\theta}_{\equiv L}\underbrace{-(1-b)\int_{\bar{x}}^{\overline{\theta}}\frac{dQ(\theta)}{d\alpha}d\theta}_{\equiv H}$$

The sign of L. $L=-B<0$.

The sign of H.

$$H=-(1-b)\Big(\int_{\bar{x}}^{\overline{\theta}}\frac{(1-U)(1-S(\theta))v}{(1-\alpha+\alpha v-U\alpha v)^2}-\frac{\partial v}{\partial\alpha}\int_{\bar{x}}^{\overline{\theta}}\frac{(1-U)(1-S(\theta))\alpha(1-\alpha)}{(1-\alpha+\alpha v-U\alpha v)^2}\Big)$$

We have already shown that $\frac{\partial v}{\partial\alpha}<0$, so $H<0$.

The sign of J. We can rewrite the expression as follows :

$$J=\frac{1}{F(\bar{x})^2}\Big(\int_0^{\underline{x}}((\underline{x}-\theta)F(\bar{x})-v)\frac{df(\theta)}{d\alpha}d\theta - v\int_{\underline{x}}^{\bar{x}}\frac{df(\theta)}{d\alpha}d\theta\Big)$$

Note that $\int_0^{\underline{x}}((\underline{x}-\theta)F(\bar{x})-v)\frac{df(\theta)}{d\alpha}d\theta$, is similar to $-A$, thus by applying the same kind of analysis we can show that it is negative. The last term :

$$-\int_{\underline{x}}^{\bar{x}}\frac{Us(\theta)v^2}{(1-\alpha+\alpha v-U\alpha v)^2}d\theta-\frac{\partial v}{\partial\alpha}\int_{\underline{x}}^{\bar{x}}\frac{Us(\theta)(1-\alpha)\alpha v}{(1-\alpha+\alpha v-U\alpha v)^2}d\theta=$$
$$-\int_{\underline{x}}^{\bar{x}}\frac{Us(\theta)v}{(1-\alpha+\alpha v-U\alpha v)^2}d\theta\Big(v+\frac{\partial v}{\partial\alpha}\alpha(1-\alpha)\Big)$$

We show that $\Big(v+\frac{\partial v}{\partial\alpha}\alpha(1-\alpha)\Big)$ is positif. Indeed :
$\Big(v+\frac{\partial v}{\partial\alpha}\alpha(1-\alpha)\Big)\geq\int_0^{\underline{x}}(\underline{x}-\theta)f(\theta)d\theta-\int_0^{\underline{x}}\frac{(\underline{x}-\theta)f(\theta)U\alpha(\underline{x}-\theta-v)}{(1-\alpha+\alpha v)+U\alpha(\underline{x}-\theta-v)}d\theta>0.$
So $J<0$.

The left hand side of the constraint decreases with α.

8.7 Characteristics of the equilibrium contract

An agent who fails to perform a task $x > \bar{x}$ keeps his occupation. Let us assume that the principal proposes a contract with efficiency wage and bonus, and that he announces that agents who fail to perform a task $x \in [\bar{x}, \bar{x} + \Delta]$, will be fired. In this case the inter-temporal expected utility of an agent $\theta < \bar{x}$ writes :

$$V(\theta) = w - \theta c + \delta(1-\alpha)V_U(\theta) + \delta\alpha[(1 - \underline{x} - \Delta + \theta)V(\theta) + (\underline{x} + \Delta - \theta)V_U(\theta)]$$

The introduction of the bonus and the threat for an agent to be fired if not performing tasks larger than $\bar{x}$, reduces the inter-temporal utility of any employee with $\theta < \bar{x}$. Thus the incentive compatible wage for problems $x \leq \underline{x}$ should be :

$$w = c(\underline{x} + \Delta) + \frac{c(1 - \alpha\delta)}{\alpha\delta} + (1 - \delta)V_U$$

We notice that if the principal increases the efficiency wage by Δ, then the payment of the bonus duplicates[26] the incentives and is not optimal.

[26] Recall that any task, independently of its difficulty, provides the same profit to the principal if solved.

9 Notations

c - effort cost

δ - discount factor

z - unemployment benefit

$(1-\alpha)$ - probability of exogenous interruption of the employer-employee relationship

N - the size of agents' population

θ - agent's ability

$S(\theta)$ - initial distribution of talent

$F(\theta)$ - ability distribution of employees

$Q(\theta)$ - ability distribution of unemployed

$\underline{x}$ - implicit performance standard for the efficiency wage

$\overline{x}$ - implicit performance standard for the bonus

v - proportion of fired employees

T - proportion of transmitted, to the principal, problems

n - number of employees

λ - the reemployment probability of an unemployed agent

V - inter-temporal expected utility of an employed agent

V_U - inter-temporal expected utility of an unemployed agent

w - fixed wage

b - bonus

U - unemployment rate

Chapitre 4

Mutual Monitoring Versus Incentive Pay in Teams

1 Introduction

Since the early nineties widespread transformations of work organization have been observed. In particular, many firms have introduced teamwork, accompanied by employee involvement practices. In a study based on 875 American firms, Osterman (1995) points out increasing adoption of self managed teams, quality programs, and job rotation.

Traditionally, the adoption of team-work results from the trade-off between a higher productivity and a higher cost of providing incentives. Indeed, when agents work together on the realization of a common project it could be source of productivity gains. For example, employees could be better informed about the best way to organize the productive process. At the other side, when agents are protected by limited liability, incentives based on team production lead to free riding (Holmstrom, 1982). In our model agents work on the realization of a common production and we focus on the possibility for the principal to reduce the cost of incentives by allowing employees to monitor each other, even if mutual-monitoring is costly.

Observations of the way firms organize their production and empirical studies provide evidence about the fact that the reward can be complemented by other management practices to improve employees' performance. Appelbaum and Batt (1994) suggest that team's productivity could be improved by ensuring employees' participation in setting human resource policies : self managed teams set "disciplinary rules governing appropriate behavior on the job and help in the selection of new entrants to the team," they could even be responsible for "developing and administering policies regarding absenteeism and the replacement of absent workers." Furthermore, recent empirical studies emphasize the existence of strong complementarities in the adoption of certain human resource practices. Ichniowski, Prennushi and Shaw (1997) as well as Boning, Ichniowski and Shaw (2003) point out that the adoption of team work improves productivity when it is accompanied by frequent interactions between team members, empowerment for solving day-to-day problems, improvements in the mutual monitoring between employees, and group remuneration schemes.

Finally the idea that when agents work in teams, monitoring is implicitly delegated to team members is also related by sociologists. Smith (1997), for example, presents the following conclusions on self-managed teams :

> "Team-based production methods represent a new, more decentered, and less visible tactic of control. Monitoring, evaluation, and disciplinary action moves down the hierarchy from the hands of supervisors and diffuses into the hands of team mates."

To shed some light on the mutual monitoring effect, we consider a model in which employees observe signals about each other's individual efforts and can contract on these signals. Employees' side-contract states a transfer and a punishment scheme. Thus if the signal about agent's effort is low, he has to pay a transfer to his colleague.

In case of low team performance and individual signal realization the agent is punished, both by the principal and his teammate. We consider that employees are subject to limited liability. If the overall transfer is too high, an agent may decide to quit the firm without paying it. Thus the reward scheme proposed by the principal affects the acceptability of the side-contract and the set of credible transfers. To allow agents to commit on a punishment transfer the employer must relax the limited liability constraint by increasing the wage paid in case of low team output realization. Then the principal's benefit from agents' side-contracting is a reduction of the incentive bonus, and the cost is an increase of the transfer paid in case of low team output realization. Hence, allowing agents to side-contract is in the principal's interest only if they are sufficiently well informed about each other's effort.

The form of the contract also depends on agent's liability limit. If the employment relationship is sufficiently valuable for an agent (*i.e.* slack limited liability constraint) the principal offers a contract with a higher bonus and use harsher punishment in case of failure. In this case the intensity of incentives provided through the side-contract is weaker.

Finally, if supervision requires costly effort form the agents, it could be in the principal's interest to delegate this task to one of the team members. We compare two possible ways for organizing supervision inside the team : mutual monitoring, both agents spend the monitoring effort and side-contract on the observable information and unilateral supervision, one of the agents supervises and proposes a contract to his colleague. We show that mutual supervision is more profitable for higher costs of the supervision effort.

In the existing literature there are two main ways to consider coordination between agents. First, cooperation can be viewed as the possibility for an agent to help his colleague in accomplishing a task, the relevant questions being that of the corresponding incentive scheme and of the optimal choice of task clustering between employees (see Itoh (1991), Macho-Stadler and Perez-Castrillo (1993) among others). A second strand of the literature, closer to our work, considers cooperation through the agents' possibility to side-contract on their action choices. Holsmtröm and Milgrom (1990) and Itoh (1993) show that the principal benefits from letting the agents side-contract on their effort choices when they can perfectly observe each other's effort[1]. Since agents can monitor each other they coordinate their effort choice. Thus the principal imposes less risk to the employees, without weakening incentives. In our model we emphasize the effects of imperfect monitoring technology and limited liability on side-contracting and on the incentive contract proposed by the principal. Another novel point of the chapter is the discussion of the endogenous adoption of monitoring technology and in particular the comparative analysis of mutual monitoring and unilateral supervision.

Finally, our paper is related to the work of Kandel and Lazear (1992). They explicitly consider the disutility effect of peer pressure without addressing the question of its endogenous formation.

The outline of this chapter is as follows. In Section 2 we present the framework and a benchmark of individual incentive scheme. In Section 3 we present the coordination agreement and derive the characteristics of the optimal mutual monitoring incentive pay mix. In Section 4 we compare mutual supervision to unilateral delegation. Section 5 concludes. All proofs are in the appendix.

2 The Model

2.1 Framework

An employer (principal) contracts with two identical employees (agents) for the realization of a project. All parties are risk neutral. However we assume that the

[1] Arnott and Stiglitz (1991) and Laffont and Rey (2001) provide similar results in the cases of insurance and micro-finance, respectively.

agents are subject to limited liability : the net transfers they receive must always be greater than or equal to some exogenous level $(-M)$, where $M \geq 0$. Each agent $i \in \{1, 2\}$ exerts a costly effort e_i that can take two possible values normalized as follows : $e_i = 0$ if the agent "shirks" and $e_i = 1$ if he "works". Exerting effort is a source of disutility for the employees : $c(0) = 0$ and $c(1) = c > 0$. The principal observes the total production level y, which is either high y^H or low y^L. As total output is the unique contractible variable, payments can only be contingent on its realization : $w_i = \{w_i(y^H); w_i(y^L)\} = \{w_i{}^H; w_i{}^L\}$. $\Delta w = w^H - w^L$ denotes the (incentive) bonus. The stochastic influence of effort on y is characterized by the probability function : $Pr(y = y^H / e_1, e_2) = p_{e_1 e_2}$.

Assumption 4.1. $1 > p_{11} > p_{01} = p_{10} > p_{00}$

The probability of high output is increasing in effort. For the sake of expositional simplicity, and without loss of generality, we normalize $p_{00} \equiv 0$.

Assumption 4.2. $p_{11} - p_{01} > p_{10} - p_{00}$

We assume a strategic complementarity[2] in employees' efforts, an agent's work raises his teammate's gain from working.

Finally, we consider that a higher effort realization is sufficiently valuable for the principal, so that it is always in his interest to provide incentives for effort provision. Thus we focus our attention on the incentive cost of implementing high effort by both agents.

2.2 Benchmark : Individually incentive contracts

In this section we present a useful benchmark of individual incentive contracts, the unique source of incentives for the agents is the wage scheme proposed by the

[2] With this assumption we rule out collective deviations from the $(1, 1)$ equilibrium.

principal to each of them. The employer's program writes as follows :

$$\min_{w^H, w^L} C = 2(p_{11}w^H + (1-p_{11})w^L) \Leftrightarrow \min_{\Delta w, w^L} C = 2(p_{11}\Delta w + w^L)$$

subject to :

$$\begin{cases} p_{11}\Delta w + w^L - c \geq p_{01}\Delta w + w^L & (IC) \\ p_{11}\Delta w + w^L - c \geq \bar{U} & (IR) \\ w^L \geq -M, w^H \geq -M & (LL) \end{cases}$$

(IC) guarantees employee's individual incentives to work. (IR) ensures agent's participation to the contract. Thereafter we normalize the outside option ($\bar{U}$) to 0. Finally, transfers must satisfy agent's limited liability constraint

The optimal contract[3] resulting from the resolution of this program is :

- $\Delta w = \dfrac{c}{p_{11} - p_{01}}$
- $w^L = \max\{-M; c - p_{11}\Delta w\}$

itemize

If the limited liability constraint is sufficiently slack ($M \geq \dfrac{cp_{01}}{p_{11} - p_{01}}$) the principal can implement the first best (then $C = 2c$). We denote $M_0 = \dfrac{cp_{01}}{p_{11} - p_{01}}$. For the rest of the chapter we focus on cases where the first best is not implementable.

Assumption 4.3. *We assume $M < M_0$.*

Since individual compensation is based on group output and agents are subject to limited liability, there is a free riding problem, which raises the incentive cost for the principal. Indeed the bonus used to motivate an agent to work, simply confers a positive externality to his teammate without improving the latter's incentives to exert effort.

3 It is easy to show that with the remuneration scheme presented below and under the assumption of effort complementarity, there are two possible equilibria : $(1;1)$ and $(0;0)$. The equilibrium selection not being the main point of our work we simply focus on the Pareto dominant equilibrium in which both agents provide high effort.

Agent's limited liability plays an important role in our setting, thus we precise our interpretation of this assumption. If the transfer (or sum of transfers) an agent has to pay exceeds M, he may decide to quit[4] the firm without paying it (or them). Considering a fixed M that does not depend on the contract proposed by the principal is clearly a short cut. However it allows us, in a tractable way, to point out the interaction between the contract proposed by the employer and the one "signed" by the co-workers.

3 Mutual monitoring

We first characterize the side-contract of team members. Then we derive the optimal incentive contract proposed by the principal under the possibility for agents to side-contract.

3.1 Presentation

Proximity between agents working together, close technological relation between their tasks, job rotation, these characteristics of team work, allow us to consider that employees could be better informed about their colleagues' actions than the employer.

In particular, we assume that once an agent has provided effort or not, all team members observe a signal $s_i = \{0;1\}$, contingent on his effort. The signal is only observable in the team, and cannot be seen by the principal. We denote the corresponding probabilities of realization as follows : $Pr(s = 0/e = 0) = q$ and $Pr(s = 0/e = 1) = r$, with $q > r$. Thus $s = 0$ is a "bad" signal if agents decide to coordinate on high effort.

[4] When an agent quits the organization he looses his expected benefits of the ongoing relationship. Levin (2003) points out that in dynamic relational contracts self-enforcing constraints act like limited liability constraints. Indeed self enforceability imposes a lower and upper bound on the transfers stated by the contract.

Agents can write an enforceable[5] side-contract on any observable (signals and output). The terms of this contract specify an effort pair, a transfer t, and a punishment scheme that maximize the joint utility function of the team members.

In our analysis t is considered as the money value of a possibly non monetary transfer. Thus there may be a gap between the value attached to the transfer by the donor and the one attached to it by the recipient[6] : we assume that a transfer t, paid by an agent to his team-mate, generates a private benefit λt for the latter[7], with $\lambda \in [0; 1]$. The imperfection of the coordination technology (when $\lambda < 1$) coupled to the possibility to observe a "bad" signal even if the agent has chosen a high effort generates a cost of side-contracting.

Timing of the game :

t=0 The principal proposes a contract to the agents $(\Delta w; w^L)$.

t=1 Each agent accepts or refuses. If one of them refuses, the game ends.

t=2 The agents decide whether or not to sign a side-contract. If one of them refuses, both are acting non cooperatively, according to the terms of the contract proposed by the principal.

t=3 Efforts are chosen, the output and the signals observed and the contract(s) executed.

Characteristics of the side-contract : There is no *ex ante* restriction on the characteristics of the side-contract, in the sense that transfers can be contingent on any information, observable by all employees. However the contract proposed by

---

5 The enforcement of side-contracts is relying on non judicial mechanisms, such as reputational devices, dynamic relations with possible trigger strategies. In this paper we assume enforceability, Che and Yoo (2001) address the question of endogenous enforcement of the side-contract.

6 A complete discussion on side-transfer technologies is provided in Tirole (1992)

7 Another way to consider the punishment is that in case of bad signal realization an agent is punished by his colleague. The disutility of a punished agent is t, and the cost of punishing someone is βt, with $\beta < 1$.

the principal limits the set of transfers on which agents can credibly commit, thus affecting the form of the feasible side-contracts. In Appendix 7.1 we show that :

Lemma 4.1. *The side-contract adopted by the employees is such that each agent pays the transfer (t) in case of "bad" signal realization, independently of the other agent's signal or the team output.*

We apply the result of Lemma 4.1, thus the coordination agreement on $(1;1)$ has to maximize the agents' joint utility[8] under the following constraints :

$$\begin{cases} p_{11}\Delta w + w^L - c - r(1-\lambda)t \geq p_{00}\Delta w + w^L & \text{(CIR)} \\ p_{11}\Delta w + w^L - c - r(1-\lambda)t \geq p_{01}\Delta w + w^L - (q-\lambda r)t & \text{(CIC)} \\ w^L - t \geq -M & \text{(CLL)} \end{cases}$$

Let us comment on each of these inequalities. First, the participation constraint (CIR) : an agent accepts the side-contract only if the expected utility of doing so exceeds his utility of acting non cooperatively in conformity with the principal's contract. There is room for coordination through side-contracting only if $\frac{c}{p_{11}} \leq \Delta w < \frac{c}{p_{11}-p_{01}}$. Indeed if $\Delta w \geq \frac{c}{p_{11}-p_{01}}$, the contract is individually incentive compatible and each agent exerts effort. Conversely, if $\Delta w \leq \frac{c}{p_{11}}$, the individual and collective choices correspond to $(0;0)$. For intermediate values of the bonus coordination through a side-contract could be valuable for the employees since the jointly profitable equilibrium is $(1;1)$, whereas the unique[9] non cooperative equilibrium is $(0;0)$. The agent's utility, in the latter case, corresponds to his outside option if the coordination agreement fails.

Note that if coordination is costless ($\lambda = 0$), the high effort equilibrium is collectively optimal (*i.e.* there is no room for simultaneous deviation) for any

[8] As the principal's objective is to implement high effort for both agents, we directly consider the characteristics of a coordination agreement on $(1;1)$. Furthermore we show in Appendix 7.2 that coordination on $(1;1)$ is the best the agents can do.

[9] Under the assumption of effort complementarity.

$\Delta w \geq \frac{c}{p_{11}}$. Here the imperfection of side-contracting technology limits the set of Δw for which coordination on $(1;1)$ occurs.

Second, incentive compatibility constraint (CIC) : since agents are side-contracting on signals, imperfectly correlated with one's effort, there is a moral hazard problem inside the team. For a given value of Δw, the punishment stated in the side-contract should be sufficiently high to prevent unilateral deviation.

Finally, transfers are constrained by agent's limited liability. As mentioned above we interpret t as the monetary value of the agent's punishment. In the worst possible situation an agent is punished at the same time by the principal (receives w^L, possibly negative) and by his teammate (pays t). The limited liability constraint states that the maximal total punishment for an agent must not exceed $(-M)$. If the constraint is not satisfied the employee prefers to quit the firm rather than to incur the corresponding costs. In order to allow agents to commit on some $t > 0$, the principal has to relax the limited liability constraint, by raising w^L. This increase in w^L is the cost the employer pays for the potential benefits of the employees' coordination.

3.2 Incentive pay when mutual monitoring is possible

The side-contracting stage :

As the cost of side-contracting is proportional to t, agents choose the minimal transfer that guarantees effort provision. From (CIC), we have[10] $t = \dfrac{c - (p_{11} - p_{01})\Delta w}{q - r}$.
This equation illustrates the substitutability between Δw and t. Since coordination is possible, an agent chooses the high effort for two reasons : receive the bonus Δw and avoid the teammate's punishment $(-t)$. For a lower Δw, employees should agree on a higher t in order to motivate each other.

However, the value of t is bounded above. First, raising t increases the cost of coordination and makes the side-contract less attractive for the agents (*i.e.* makes

[10] Recall that we made the following normalization $p_{00} \equiv 0$

(CIR) tighter), so $t \leq \dfrac{p_{11}\Delta w - c}{r(1-\lambda)}$. Second, for the punishment to be credible it should be such that $t \leq M + w^L$ (*i.e.* (CLL) holds).

The employer's contract :

Anticipating the possibility for agents to side-contract, the principal shapes the contract in order to minimize the cost of implementing the high effort equilibrium.

$$\min_{\Delta w, w^L} C = 2(p_{11}\Delta w + w^L)$$

under the constraints

$$\begin{cases} \Delta w \geq \dfrac{c-(q-r)(M+w^L)}{p_{11}-p_{01}} & (1.1) \\ \Delta w \geq \dfrac{c(q-\lambda r)}{p_{11}(q-\lambda r)-p_{01}(r-\lambda r)} & (1.2) \\ w^L \geq -M & (1.3) \\ p_{11}\Delta w + w^L - c - (r-\lambda r)\dfrac{c-(p_{11}-p_{01})\Delta w}{(q-r)} \geq 0 & (1.4) \end{cases}$$

The existence of the side-contract allows the principal to reduce the bonus proposed to the employees. However, as t is bounded above, the reduction of Δw is limited. The right hand side of constraint (1.1) corresponds to the minimal value of the bonus for which agents can agree on a transfer t that satisfies both the incentive compatibility and limited liability constraints.
Expression (1.2) states that if the employer proposes a contract with Δw, below that value, it would be impossible for the agents to commit on side-transfers guaranteeing incentive compatibility (CIC) and acceptability (CIR) of the coordination agreement. This expression corresponds to the maximal reduction of the incentive bonus that could be achieved through employees' coordination.
Finally, equation (1.3) is the limited liability constraint and equation (1.4) guarantees agent's participation to the employment contract.

The characteristics of the optimal contract depend on the informativeness of the signals and on the liability limit of the agents.

Proposition 4.1. ***Low informativeness of the signals***
If agents are not sufficiently well informed ($q - r \leq \frac{p_{11} - p_{01}}{p_{11}}$) then the optimal contract is the individually incentive one.

The individually incentive contract, obtained in Section 2.2, is $\left(w^L = -M, \Delta w = \dfrac{c}{p_{11} - p_{01}}\right)$. The transfer between agents in this case is set to zero.

The principal can relax (1.1), by increasing w^L. When w^L is raised by one unit, it allows a reduction in the incentive bonus by $\dfrac{(q-r)}{p_{11} - p_{01}}$. This operation reduces the principal's expected wage cost only if, $\left(-\dfrac{(q-r)p_{11}}{p_{11} - p_{01}} + 1\right) < 0$. The side-contract expands the principal's set of incentive instruments. However, their use depends on their relative efficiency. The condition of Proposition 4.1 illustrates the trade-off between a harder punishment (*i.e.* lower w^L) in case of failure and a lower incentive bonus (*i.e.* lower Δw) in case of success. The possibility of coordination for agents by the means of a side-contract is sufficiently valuable only if the marginal quality of the mutual monitoring technology exceeds the marginal likelihood ratio.

Proposition 4.2. ***Highly informative signals ($q - r > \frac{p_{11} - p_{01}}{p_{11}}$) and optimal incentive mix.***
There exists $\bar{M} = \dfrac{cp_{01}}{p_{11}(q - \lambda r) - p_{01}r(1-\lambda)} < M_0$, such that :

1. *If $M \leq \bar{M}$, the binding constraints are* (1.1) *and* (1.2).
$$\Delta w = \frac{c(q - \lambda r)}{p_{11}(q - \lambda r) - p_{01}(r - \lambda r)} \text{ and } w^L = \bar{M} - M$$
2. *If $M > \bar{M}$, the binding constraints are* (1.1) *and* (1.4).
$$\Delta w > \frac{c(q - \lambda r)}{p_{11}(q - \lambda r) - p_{01}(r - \lambda r)} \text{ and } w^L < \bar{M} - M$$

According to Proposition 4.1, if $(q - r) > \dfrac{p_{11} - p_{01}}{p_{11}}$ reducing the incentive bonus by allowing agents to side-contract is profitable for the principal. Thus, he increases w^L and reduces Δw to keep (1.1) satisfied. This constraint guarantees that any increase of w^L is intended to raise the side-transfer and reduce Δw. It is always binding.

There are two upper bounds for w^L given by constraints (1.2) and (1.4). First, constraint (1.2) corresponds to the maximal reduction of the bonus that can be achieved by agents' side-contracting. So once this limit has been attained, there is no need to increase w^L further. The second limit for w^L is given by the participation to the principal's contract constraint. When the employer increases w^L there is a direct raise of employee's utility, but also two indirect negative effects : it reduces the bonus and raises the cost of side-contracting. In fine the negative effects are stronger and worker's utility decreases with w^L.

For low values of M (*i.e.* $M < \bar{M}$) the binding constraint is (1.2). The agents' initial rent is sufficiently high, thus the principal can reduce as much as possible the bonus, by allowing agents to side-contract. Then we talk about total coordination. When M is high (*i.e.* $M > \bar{M}$) then increasing w^L up to the value given by (1.2) does not ensure agents' participation. So the principal reduces w^L to bind the participation constraint (1.4). We refer to this case as partial coordination. The higher M the closer the incentive bonus to its individually incentive value. When the employment relationship is sufficiently valuable for an agent (*i.e.* higher M), it is in the principal's interest to propose more incentive contracts (higher Δw), and to use harsh punishments in case of failure (lover w^L). Results from Propositions 4.1 and 4.2 are summarized on Figure 4.1.

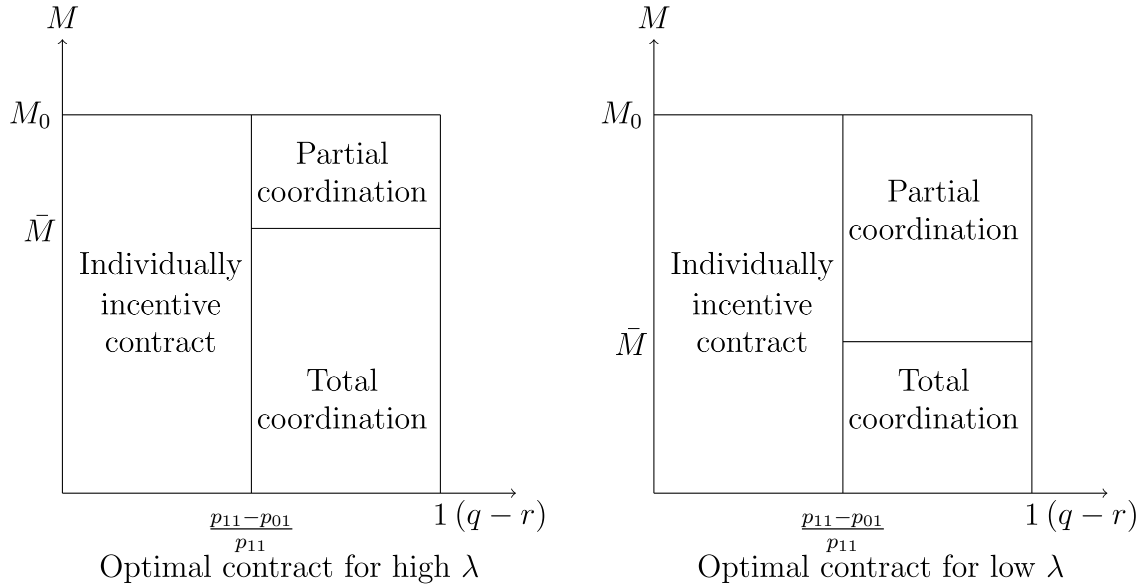


Figure 4.1 : The optimal incentive contract when employees monitor each other.

The imperfection of the side-contracting technology does not alter the relative efficiency of the informal agreement. Indeed the principal's decision to allow or not the side-contract depends only on the quality of employees' information. However λ affects the set of bonuses for which coordination is possible. A higher dead-weight loss (low λ), raises the cost of side-contracting, and thus reduces the set of Δw for which coordination is acceptable for the employees. Furthermore, for given value of the liability limit, the structure of the employment contract depends on λ. When λ is lower (high dead weight loss), $\bar{M}$ decreases, which expands the set of M for which partial, rather than, total coordination is adopted. Then increasing λ (lower dead weight loss) raises the principal's profit.

Side-contracting is profitable if employees are sufficiently well informed about each other's actions. Furthermore better information (*i.e.* higher $(q-r)$) reduces the employer's wage cost. Then it may be profitable for the principal to induce agents to supervise each other even if there is a cost of supervision. However, under costly monitoring the principal could be better off by inducing unilateral, rather than mutual supervision. In the next section we compare those two organizational choices.

4 Mutual supervision vs. Unilateral delegation

The employer can implement costly procedures (such as regular team meetings, adoption of supervision technology and so on.), that facilitate mutual supervision between agents. At least part of the cost is supported by the team members - to prepare a meeting, to learn the utilization of a new supervisory technology and so on. In this section we consider that in order to observe a signal about the other's effort an employee should spend some monitoring cost a. The supervision effort is not observable for the principal. As previously $Pr(s = 0/e = 0) = q$ and we set the probability to observe a bad signal if the other works to 0 (*i.e.* $r = 0$).

Before proposing the contract the principal chooses between two organizations, mutual supervision (MS hereafter) and unilateral delegation (DS hereafter). MS means that both agents have access to the monitoring technology, and the principal contracts with each employee. In the case of DS, the employer delegates to one of the agents the access to the monitoring technology and the contractual relationship with the other agent. We characterize the contract proposed by a principal in each of these situations. Then we discuss the principal's choice between MS and DS.

4.1 Mutual supervision (MS)

Both agents have access to the monitoring technology.

Timing :

t=0 The principal proposes the contract $(\Delta w; w^L)$ to the agents, and decides whether or not to give them access to a mutual monitoring technology.

t=1 Each agent accepts or refuses. If one of them refuses, the game ends.

t=2 Agents either sign or not a side-contract.

t=3 Each of them decides to invest or not in a supervision technology.

t=4 Productive efforts are chosen[11]

t=4 The output and the signals are observed and the contract(s) executed.

As previously, employees are maximizing their joint utility under the participation to the coordination agreement, the incentive compatibility, and limited liability constraints.

$$\begin{cases} p_{11}\Delta w + w^L - c - a \geq p_{00}\Delta w + w^L & (CIR) \\ p_{11}\Delta w + w^L - c - a \geq p_{01}\Delta w + w^L - a - tq & (CIC) \\ w^L - t \geq -M & (LL) \end{cases}$$

First, note that since supervision is realized and observed before the choice of productive effort, if the coordination agreement is acceptable for the agents, then the supervisory effort is obviously done. Second, as the transfer does not influence the agents' expected utility, on the equilibrium they choose the highest possible t. The limited liability constraint binds.

The principal‘s program writes :

$$\min_{\Delta w, w^L} C = 2(p_{11}\Delta w + w^L)$$

under the constraints :

$$\begin{cases} p_{11}\Delta w - c - a \geq 0 & (1.5) \\ (p_{11} - p_{01})\Delta w - c + (M + w^L)q \geq 0 & (1.6) \\ w^L \geq -M & (1.7) \\ p_{11}\Delta w + w^L - c - a \geq 0 & (1.8) \end{cases}$$

[11] Each agent chooses his productive effort after observing the decision of his team-mate at each of the previous stages. Thus we eliminate a possible commitment problem that arises when investment in supervisory technology is made *ex post.*

Equation (1.6) and (1.8) correspond respectively to the incentive compatibility and participation to the team contract constraints. Expression (1.5) ensures agents' participation to the side-contract. Note that, under the assumption of effort complementarity, it guarantees that simultaneous deviation from $(1,1)$ is not profitable for the agents[12].

The results are very close to those in Section (3.2).

Proposition 4.3. *The team contract* $\Delta w = \dfrac{c+a}{p_{11}}$, $w^L = \max\{-M + \dfrac{p_{01}c - a(p_{11} - p_{01})}{p_{11}q}; 0\}$ *is optimal if and only if the monitoring technology is :*

- *Sufficiently informative :* $q \geq \dfrac{p_{11} - p_{01}}{p_{11}}$
- *Not too costly :* $a \leq \hat{a}$, *where* $\hat{a} = M_0 - M$

The higher the value of M, the lower the $\hat{a}$. Thus for high values of the liability limit, the monitoring technology is adopted only if its cost is very low. This observation is consistent with the result of Proposition 4.2. Indeed in Section 3.2, the cost of coordination is related to the imperfection of the coordination technology, and is proportional to the transfer t. Thus for high values of M, the principal may want to reduce (implicitly) the coordination cost by proposing a contract with larger incentive bonus. Now there is a fixed cost of coordination, thus the question is not how much mutual monitoring is optimal for the principal, but whether it is optimal or not. The adoption of mutual monitoring makes sense for the employer only if he is strongly constrained in the possibility to punish agents in case of low output.

4.2 Unilateral delegation (DS)

The question of optimal delegation in an organization subject to moral hazard is investigated by Baliga and Sjostrom (1998). They characterize the optimal delegation in a model where one of the agents observes without cost the other's

[12] The complete constraint for $(1;1)$ to be the couple of efforts maximizing the agents' joint utility is : $2(p_{11}\Delta w + w^L - c - a) \geq \max\{2(p_{00}\Delta w + w^L); 2(p_{01}\Delta w + w^L) - c - a\}$. It is easy to show that under the assumption of effort complementarity it is equivalent to $p_{11}\Delta w - c - a \geq 0$

effort. If employees are symmetric, as in our setting, it is optimal to delegate to the informed employee the capacity to contract with his colleague. Our paper compares, from the principal's point of view, unilateral and bilateral monitoring, when the latter is costly. The principal is not allowed to use the monitoring technology[13] himself. Teammates already have some advantage in supervising each other and the adoption of monitoring technology is less costly for them than for the principal.

Let us consider, without loss of generality since agents are identical, that in case of DS the principal delegates monitoring and contracting to A_1. We set $M \equiv 0$.

Timing in the case of delegation :

t=0 The principal proposes the contract $(\Delta w; w^L)$ to A_1, and gives him access to a supervisory technology $(a; q)$.

t=1 A_1 accepts or refuses. If he refuses the game ends.

t=2 If A_1 accepts, he proposes a contract $(t^H; t^L)$ to A_2.

t=3 A_2 accepts or refuses.

t=4 Efforts are chosen[14], and contracts executed.

The contract the principal proposes to A_1 should motivate him to exert both the supervisory and the productive effort. It also should ensure that A_1 prefers to motivate A_2 rather than to be the only working agent and keep the total bonus if the projects succeeds. A detailed presentation of the principal's program under unilateral delegation is proposed in Appendix 7.5.

Proposition 4.4. ***Profitable delegation :*** *There exists*[15] $\bar{a} \leq \dfrac{cp_{01}}{p_{11} - p_{01}}$, *such that for* $a \leq \bar{a}$, *it is valuable for the principal to delegate contracting to the supervisor.*

[13] Demougin and Fluet (2001) study the trade-off between monitoring and incentives in a principal-agent model with moral hazard.

[14] To preserve the consistency with the case of mutual supervision, we consider that before choosing his productive effort A_2 observes if the supervisory effort has been done by A_1.

[15] $\bar{a} = \dfrac{cp_{01}p_{11}q}{(p_{11} - p_{01})(p_{11} - p_{01}(1 - q))}$

The principal cannot observe if the supervision effort has been made or not, thus he cannot propose to A_1 a contract contingent on it. Supervising the agent A_2, allows A_1 to reduce his bonus. Thus A_1 spends a only if it does not exceed the gain from the reduction of the bonus. When a is larger than $\bar{a}$, A_1 never spends the cost of supervision, thus the optimal contract with decentralization corresponds to the individually incentive compatible centralized[16] contract.

Proposition 4.5. ***Optimal contract with delegation :*** *There exists*[17] $\tilde{a} < \bar{a}$ *such that :*

1. *If* $a \leq \tilde{a}$ *the contract proposed to* A_1 *is :*
   $$\underline{\Delta w} = \frac{c}{p_{11} - p_{01}} + \frac{c}{p_{11} - (1-q)p_{01}}$$
   $$w^L = 0.$$
2. *If* $\tilde{a} < a \leq \bar{a}$ *the contract proposed to* A_1 *is :*
   $$\overline{\Delta w} = \underline{\Delta w} + \frac{a}{p_{11} - p_{01}} - \frac{c{p_{01}}^2 q}{(p_{11} - p_{01})^2 (p_{11} - p_{01}(1-q))}$$
   $$w^L = 0.$$

For low a ($a < \tilde{a}$), supervision is valuable for A_1, independently of his productive effort. The principal's contract motivates him to work. We notice that the wage is independent of the cost of supervision. When $a > \tilde{a}$, supervision is valuable only if the supervisor exerts high effort. In this case the incentive bonus should ensure that A_1 makes both efforts. It is increasing in a.

4.3 MS vs. DS

We now derive the conditions under which the principal chooses each of these contractual forms.

Proposition 4.6. ***MS vs. DS :*** *There exists* $\bar{\bar{a}} \in [\tilde{a}, \bar{a}]$*, such that :*

[16] The principal contracts with each agent.

[17] $\tilde{a} = \dfrac{c{p_{01}}^2 q}{(p_{11} - p_{01})(p_{11} - p_{01}(1-q))}$

– *For* $a \leq \bar{\bar{a}}$*, DS dominates MS.*

– *For* $\bar{\bar{a}} < a \leq \frac{cp_{01}}{p_{11} - p_{01}}$*, MS dominates DS.*

The proposition states that for higher costs of the monitoring technology MS is more valuable. This result may seem surprising at first sight since under MS both agents spend the cost a.
In the case of DS, the A_1's wage is higher than the one of both workers under MS, which in turn is higher than the wage of A_2. For technologies with low supervision cost, the latter effect is stronger than the former, so it is in the principal's interest to adopt DS. For costly monitoring technology the bonus of the supervising agent has to provide incentives both to work and monitor. Thus we attain some value a above which MS becomes more profitable for the principal.

5 Conclusion

Organizational theorists have emphasized the multilateral nature of contracting in firms, thus suggesting that they could be considered as nexus of contracts. Indeed, any collegial or hierarchical relationship in the organization involves some (side-)contractual form. In this chapter we examine the interactions between the contract proposed by the principal and the possibility for employees to side-contract and monitor each other. When agents side-contract, each employee works both to earn the reward proposed by the principal and to avoid the punishment by a teammate. We assumed that agent's tolerance for punishments, from the principal or from a teammate, is limited. Employees can commit on a credible punishment only if the principal relaxes the limited liability constraint and raises the wage paid when employees' production is low. The principal has to choose between a lower incentive bonus and a higher punishment. Thus we show that agents' side-contracting is profitable for the principal since team members are sufficiently well informed about each other's performance. The reduction of the bonus is more important when agents limit liability is more stringent. Finally, we discuss the conditions for endogenous adoption of mutual monitoring technology and compare unilateral and bilateral supervision from the principal's point of view.

The efficiency of mutual monitoring is an argument for keeping work teams small (Roberts (2004)). In our setting we set aside considerations of the optimal team size. Thus an interesting extension of the model could be to investigate endogenous team formation.

Another limitation of this paper that could be addressed in a future work is in providing a more complete analysis of the interaction between contracts conditionally on their more or less enforceable character. Relational contracts on both sides, formal contract proposed by the principal and relational side-contract and so on.

6 References

Appelbaum E. and R. Batt (1994), *The New American Workplace*, Cornell University Press.

Arnott R. and J. Stiglitz (1991), "Moral Hazard and Non-Market Institutions : Dysfunctional Crowding Out of Peer Monitoring", *American Economic Review*, vol. 81, pp. 179-190.

Baliga S. and T. Sjostrom (1998), "Decentralization and Collusion", *Journal of Economic Theory*, vol. 83, pp. 196-232.

Boning W., C. Ichniowski and K. Show (2003), "Opportunity Counts : Teams and the Effectiveness of Production Incentives", NBER Working Paper.

Che Y-K. and S-W. Yoo (2001), "Optimal Incentives in Teams", *American Economic Review*, vol. 81, pp. 525-541.

Demougin D. and C. Fluet (2001), "Monitoring versus Incentives", *European Economic Review*, vol. 45, pp. 489-496.

Holmström B. (1982), "Moral Hazard in Teams", *The Bell Journal of Economics*, vol. 13, pp. 324-340.

Holmström B. and R. Milgrom (1990), "Regulating Trade Among Agents", *Journal of Institutional and Theoretical Economics*, vol. 146, pp. 85-105.

Ichniowski C., G. Prennushi and K. Shaw (1997), "The Effects of Human ressources Management Practices on Productivity", *American economic Review*, vol. 59, pp. 291-313.

Itoh H. (1991), "Incentives to Help in Multi-agents Situations", *Econometrica*, vol. 59, pp. 611-636.

Itoh H. (1993), "Coalitions, Incentives and Risk Sharing", *Journal of Economic Theory*, vol. 60, pp. 410-427.

Kandel E. and E. Lazear (1992), "Peer Pressure and Partnerships", *The Journal of Political Economy*, vol. 100, pp. 801-817.

Laffont JJ. and P. Rey (2001), "Collusion and Group Lending with Moral Hazard", *IDEI Working Paper*.

Levin J. (2003), "Relational Incentive Contracts", *American economic Review*, vol. 93, pp. 835-847.

Macho-Stadler I. and D. Perez-Castrillo (1993), "Moral Hazard with Several Agents, the Gains from Cooperation", *International Journal of Industrial Organization*, Vol. 11, pp. 73-100.

Osterman P. (1994), "How Common is Workplace Transformation and Who adopts It ?", *Industrial and Labor Relations Review*, Vol. 47(2), pp. 173-188.

Roberts J. (2004), "The Modern Firm : Organizational Design for Performance and Growth", *Oxford University Press*.

Smith V. (1997), "New Forms of Organization", *Annual Review of Sociology*, Vol. 23, pp. 315-339.

Tirole J. (1992), "Collusion and the Theory of Organizations", *Advances in Economic Theory Sixth World Congress*, vol. 2, pp. 151-207.

7 Appendix

7.1 Proof of Lemma 4.1

The coordination agreement on $(1,1)$ has to satisfy the following constraints :

$$\begin{cases} p_{11}\Delta w + w^L - c - \underline{\alpha} t \geq p_{00}\Delta w + w^L & (CIR) \\ p_{11}\Delta w + w^L - c - \underline{\alpha} t \geq p_{01}\Delta w + w^L - \overline{\alpha} t & (CIC) \\ w^L - t \geq -M & (CLL) \end{cases}$$

where $\underline{\alpha}$ ($\overline{\alpha}$) corresponds to the expected probability[18] net of coordination cost for a working (shirking) agent to pay t.

To proof the result of Lemma 4.1 we proceed in three steps :

1. We identify the coordination agreement that is preferred by a principal.
2. We show that a principal can restrict the set of possible side-contracts to the scheme he prefers.
3. We find the side-contracting scheme corresponding to that criterion.

First Step For given t, $\Delta w = \dfrac{c - (\overline{\alpha} - \underline{\alpha})t}{p_{11} - p_{01}}$ from (CIC), and $w^L = t - M$ from (LL). Thus for given t the cost of the principal when agents are side-contracting writes :

$$C = 2\big(p_{11}\Delta w + w^L\big) = 2\big(p_{11}(\frac{c - (\overline{\alpha} - \underline{\alpha})t}{p_{11} - p_{01}}) + t - M\big)$$

The coordination agreement is profitable for the principal if : $(\overline{\alpha} - \underline{\alpha}) > \frac{p_{11} - p_{01}}{p_{11}}$. If this condition is satisfied it is immediate to see that the principal's cost is minimized for the largest possible $(\overline{\alpha} - \underline{\alpha})$.

Second Step Let $(\underline{\alpha}; \overline{\alpha}; t)$ correspond to the side-contract that maximizes the principal's profit, and $(\underline{\widetilde{\alpha}}; \widetilde{\overline{\alpha}}; \widetilde{t})$ to the one that maximizes the employees' joint utility.

[18] $\underline{\alpha}$ and $\overline{\alpha}$ are some linear combinations of r, q, λ, p_{11} and p_{01}, conditionally on the punishment scheme adopted by the agents.

The agents will be able to implement the high effort equilibrium with such a contract only if :

$$\widetilde{t} \leq t \Leftrightarrow \frac{c-(p_{11}-p_{01})\Delta w}{\widetilde{\overline{\alpha}}-\underline{\widetilde{\alpha}}} \leq \frac{c-(p_{11}-p_{01})\Delta w}{\overline{\alpha}-\underline{\alpha}} \Leftrightarrow (\overline{\alpha}-\underline{\alpha}) < (\tilde{\overline{\alpha}}-\underline{\tilde{\alpha}})$$

As we have seen the contract that minimizes the principal's wage cost is the one with the largest $(\overline{\alpha}-\underline{\alpha})$. Thus the agents are constrained to side-contract on the contract that minimizes the principal's wage cost.

Third Step The different punishment schemes we have to compare could be based on :

- Each agent's signal $(\underline{\alpha}(s_i);\overline{\alpha}(s_i))$, $\underline{\alpha}(s_i)=r(1-\lambda)$ and $\overline{\alpha}(s_i)=(q-\lambda r)$
- Both signals $(\underline{\alpha}(s_i;s_j);\overline{\alpha}(s_i;s_j))$, $\underline{\alpha}(s_i;s_j)=r(1-r)(1-\lambda)$ and $\overline{\alpha}(s_i;s_j)=q(1-r)-r(1-q)\lambda$
- Each agent's signal and output $(\underline{\alpha}(s_i;y);\overline{\alpha}(s_i;y))$, $\underline{\alpha}(s_i;y)=(1-p_{11})r(1-\lambda)$ and $\overline{\alpha}(s_i;y)=(1-p_{01})(q-\lambda r)$
- Both signals and output $\underline{\alpha}(s_i;s_j;y);\overline{\alpha}(s_i;s_j;y)$, $\underline{\alpha}(s_i;s_j;y)=(1-p_{11})r(1-r)(1-\lambda)$ and $\overline{\alpha}(s_i;s_j;y)=(1-p_{01})(q(1-r)-r(1-q)\lambda)$

where s_i is the agent's signal, s_j is his team mate's signal, and y is the output realization. We denote $\Delta\alpha(\cdot)\equiv\overline{\alpha}(\cdot)-\underline{\alpha}(\cdot)$. Thus we have :

$$\begin{aligned}
&\Delta\alpha(s_i)=(q-r)\\
&\Delta\alpha(s_i;s_j)=(q-r)(1-r+\lambda r)\\
&\Delta\alpha(s_i;y)=(q-r-\lambda r(p_{11}-p_{01})-p_{01}q+p_{11}r)\\
&\Delta\alpha(s_i;s_j;y)=((q-r)(1-r+\lambda r)-p_{01}q(1-r+\lambda r)+p_{01}\lambda r+p_{11}r(1-r)(1-\lambda))
\end{aligned}$$

Thus we show that under the assumption of effort complementarity $\Delta\alpha(s_i;y) > \Delta\alpha(s_i;s_j;y)$, $\Delta\alpha(s_i)>\Delta\alpha(s_i;s_j)$, and $\Delta\alpha(s_i)>\Delta\alpha(s_i;y)$.

7.2 Agents coordinate on $(1;1)$.

In $t = 3$, for given terms of the team contract $(\Delta w; w^L)$, the agents are choosing $(t; e_1; e_2)$ to maximize their joint utility.

The joint utility of the agents when they coordinate on $(0;1)$, writes.

$$U(0;1) + U(1;0)$$

For the interval of incentive bonuses we are interested on, $(0;0)$ is the unique noncooperative equilibrium. So one of the agents will work only if he receives the adequate incentives for doing so. Punishing him in case of bad signal realization is not sustainable. Thus we study the case when he receives a bonus from the other agent in case of "good" signal realization.

The following individual constraints must be satisfied :

$$\begin{cases} p_{01}\Delta w + w^L - c + (1-r)\lambda t > p_{00}\Delta w + w^L + (1-q)\lambda t & (IC1) \\ p_{01}\Delta w + w^L - (1-r)t > p_{11}\Delta w + w^L - c - (1-r)t\lambda & (IC2) \\ p_{01}\Delta w + w^L - c + (1-r)\lambda t > p_{00}\Delta w + w^L & (IR1) \\ p_{01}\Delta w + w^L - (1-r)t > p_{00}\Delta w + w^L & (IR2) \\ t + w^L \leq M & (LL) \end{cases}$$

From (IC1), we have $t \geq \dfrac{c - p_{01}\Delta w}{(q-r)\lambda}$.

The incentive compatible transfer implementing $(0;1)$ is higher than the one ensuring the agents' coordination on $(1;1)$ (under the assumption of effort complementarity). As far as the transfers the agents can commit on are constrained by the contract proposed by the principal, he will relax the limited liability constraint just enough to implement the high effort equilibrium.

7.3 Proof of Propositions 4.1 and 4.2

Agents maximize their joint utility function under the following individual constraints :

$$\begin{cases} p_{11}\Delta w + w^L - c - t(r-\lambda r) \geq p_{00}\Delta w + w^L & (CIR) \\ p_{11}\Delta w + w^L - c - t(r-\lambda r) \geq p_{01}\Delta w + w^L - c - t(q-\lambda r) & (CIC) \\ t \leq M + w^L & (CLL) \end{cases}$$

They fix the transfer at its minimal value :

$$t = \frac{c - (p_{11} - p_{01})\Delta w}{q - r}$$

t is bounded above by (CLL) and (CIR)

$$\begin{cases} t \leq M + w^L & \Leftrightarrow \quad \Delta w \geq \dfrac{c - (q-r)(M + w^L)}{p_{11} - p_{01}} \\ t \leq \dfrac{p_{11}\Delta w - c}{r(1-\lambda)} & \Leftrightarrow \quad \Delta w \geq \dfrac{(q-\lambda r)}{p_{11}(q-\lambda r) - p_{01}(r - \lambda r)} \end{cases}$$

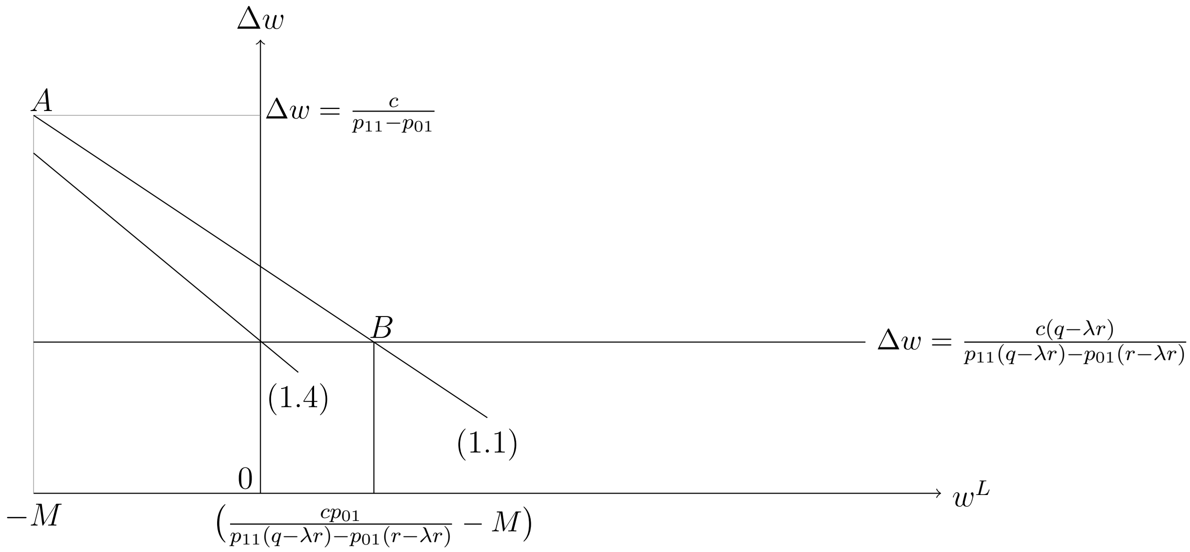
The principal's program :

$$\max_{\Delta w, w^L} -(p_{11}\Delta w + w^L)$$

under the constraints :

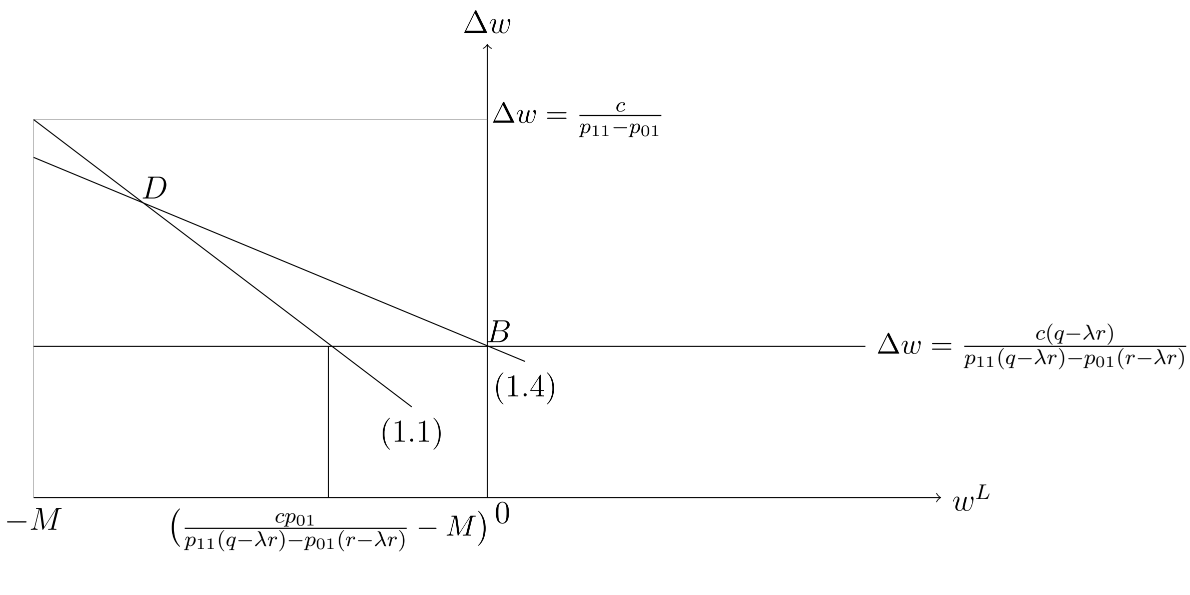
$$\begin{cases} \Delta w \geq \dfrac{c-(q-r)(M+w^L)}{p_{11}-p_{01}} & (1.1) \\ \Delta w \geq \dfrac{c(q-\lambda r)}{p_{11}(q-\lambda r) - p_{01}(r-\lambda r)} & (1.2) \\ w^L \geq -M & (1.3) \\ p_{11}\Delta w + w^L - c - (r-\lambda r)\dfrac{c - (p_{11}-p_{01})\Delta w}{(q-r)} \geq 0 & (1.4) \end{cases}$$

First case : for $M < \dfrac{cp_{01}}{p_{11}(q-\lambda r) - p_{01}(r-\lambda r)}$



There are two possible solutions for the principal corresponding to points A (individually incentive contract) and B (contract with side-contracting). As shown above, if $(q-r) > \dfrac{p_{11}-p_{01}}{p_{11}}$, the principal prefers the contract corresponding to B. For low values of the liability limit, the binding constraints are (1.1) and (1.2).

Second case : for $M > \dfrac{cp_{01}}{p_{11}(q-\lambda r)-p_{01}(r-\lambda r)}$



The incentive cost associated to the point B is :
$C_B = \dfrac{p_{11}c(q-\lambda r)}{p_{11}(q-\lambda r)-p_{01}(r-\lambda r)}$. The incentive cost associated to the point D is
$C_D = \dfrac{cp_{01} - cp_{11}(1-(q-\lambda r)) - M(p_{11}-p_{01})(r-\lambda r)}{p_{11}(q-\lambda r)-p_{01}(r-\lambda r)-(p_{11}-p_{01})}$.

After simple arithmetic transformations we show that $C_D < C_B$, for $M > \dfrac{cp_{01}}{p_{11}(q-\lambda r)-p_{01}(r-\lambda r)}$. The optimal solution binds (1.1) and (1.4).

7.4 Proof of Proposition 4.3

The program of the principal writes :

$$\max_{\Delta w, w^L} -2(p_{11}\Delta w + w^L)$$

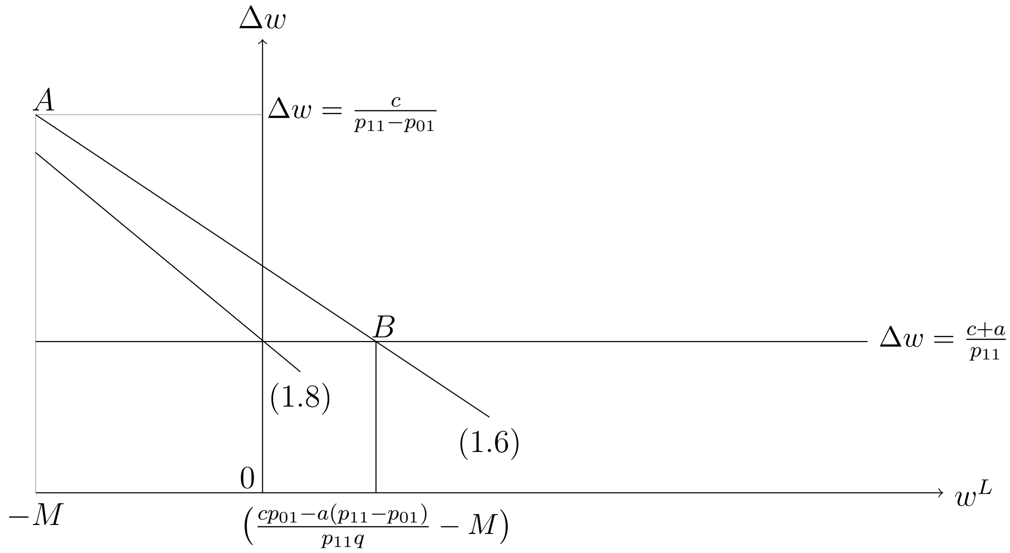
under the constraints :

$$\begin{cases} p_{11}\Delta w - c - a \geq 0 & (1.5) \\ (p_{11}-p_{01})\Delta w - c + (M+w^L)q \geq 0 & (1.6) \\ M + w^L \geq 0 & (1.7) \\ p_{11}\Delta w + w^L - c - a \geq 0 & (1.8) \end{cases}$$

– Proof of the condition $q \geq \dfrac{p_{11}-p_{01}}{p_{11}}$

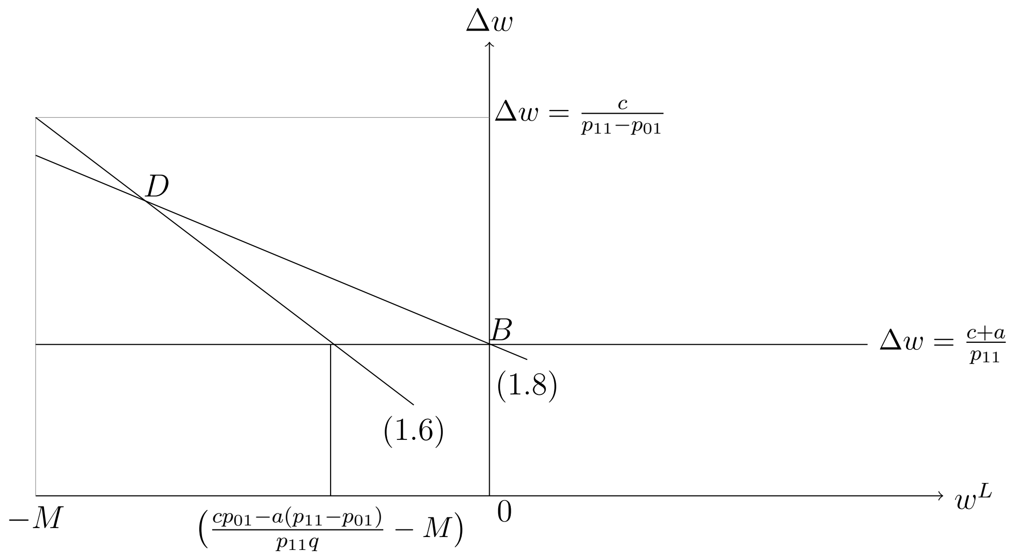
If a principal increases w^L with one unit, it makes possible a reduction of the incentive bonus by $\dfrac{q}{p_{11}-p_{01}}$. The final effect on the principal's wage cost is : $1 - \dfrac{p_{11}q}{p_{11}-p_{01}}$. The cost decreases only if $q \geq \dfrac{p_{11}-p_{01}}{p_{11}}$.

First case : for $a < \dfrac{cp_{01} - Mqp_{11}}{p_{11}-p_{01}}$



In this case and under the condition $q \geq \dfrac{p_{11}-p_{01}}{p_{11}}$ the optimal contract corresponds to B and writes : $\Delta w = \dfrac{c+a}{p_{11}}$, $w^L = \dfrac{cp_{01} - a(p_{11}-p_{01})}{p_{11}q} - M$.

Second case : for $\dfrac{cp_{01} - Mqp_{11}}{p_{11}-p_{01}} < a$



The incentive cost of the contract corresponding to the point (B), is $C_B = c + a$. The incentive cost corresponding to the contract of point (D) writes $C_D = c + a - w^L + w^L = c + a$. Thus any contract between the point (B) and (D) is optimal for the principal.

The cost of the individually incentive contract writes : $C = \frac{p_{11}c}{p_{11} - p_{01}} - M$. Then it is immediate to see that $C_D < C$ for $a < \frac{cp_{01}}{p_{11} - p_{01}} - M$.

7.5 Unilateral delegation

The principal delegates supervision and contracting to A_1. Recall that $M \equiv 0$.

– The contract proposed by A_1 to A_2.

If A_1 doesn't spend the supervision effort : $\Delta t^{NS} = \frac{c}{p_{11} - p_{01}}$ and $t_L = 0$.

If A_1 spends the supervision effort : $\Delta t^{S} = \frac{c}{p_{11} - p_{01}(1-q)}$ and $t_L = 0$.

– The principal's program.

$$\max_{\Delta w, w^L} -(p_{11}\Delta w + w^L)$$

under the constraints :

$$\begin{cases} p_{11}(\Delta w - \Delta t^S) - c - a \geq p_{11}(\Delta w - \Delta t^{NS}) - c & (1.9) \\ p_{11}(\Delta w - \Delta t^S) - c - a \geq p_{01}(\Delta w - \Delta t^S) - a & (1.10) \\ p_{11}(\Delta w - \Delta t^S) - c - a \geq p_{01}(\Delta w - \Delta t^{NS}) & (1.11) \\ p_{11}(\Delta w - \Delta t^S) - c - a \geq p_{01}\Delta w - c & (1.12) \\ p_{11}(\Delta w - \Delta t^S) + w^L - c - a \geq 0 & (1.13) \\ w^L \geq 0 & (1.14) \end{cases}$$

Constraint (1.9) ensures incentives for supervision. (1.10) is the incentive compatibility constraint for effort provision. Constraint (1.11) guarantees that A_1 does not shirk on both efforts. (1.12) guarantees that the supervisor will provide the right incentives to work to the other agent. Finally we have participation and limited liability constraints.

From constraint (1.9) we obtain the maximal value for the supervision cost $\bar{a} = \frac{cp_{11}p_{01}q}{(p_{11} - p_{01})(p_{11} - p_{01}(1-q))}$ above which it is impossible to provide incentives to A_1 to make the supervision effort. If there is no supervision - there is equivalence between delegation and centralized contracting, *i.e.* the contract proposed to the agent A_1 in this case replicates the centralized individual incentive contract.

The sign of the next expression $p_{01}(\Delta t^{NS} - \Delta t^{S}) - a \equiv B$ determines which of the constraints (1.10) or (1.11) will be binding, which will fix the Δw value.

If $B > 0$ ($\Leftrightarrow a < \dfrac{cp_{01}^2 q}{(p_{11}-p_{01})(p_{11}-p_{01}(1-q))} \equiv \tilde{a}$) the binding constraint is (1.10), and conversely for $B < 0$ ($a > \tilde{a}$), the binding constraint is (1.11).

The constraint (1.12) is always slack, under the assumption of effort complementari-ty.

7.6 Proof of Proposition 4.6

The function $G(a)$ is the difference between the cost of the principal in the case of unilateral delegation and in the case of mutual monitoring.

- If $G(a) > 0$, MS dominates DS.
- If $G(a) < 0$, DS dominates MS.

To compare those organizational structures we focus on the case MS is potentially profitable for the principal, *i.e.* $q > \dfrac{p_{11}-p_{01}}{p_{11}}$

First case : for $a < \tilde{a}$

$$G(a) = \frac{p_{11}c}{p_{11}-p_{01}} + \frac{p_{11}c}{p_{11}-(1-q)p_{01}} - 2(c+a) - \frac{2(p_{01}c-(p_{11}-p_{01})a)}{p_{11}q}$$

$$\frac{\partial G(a)}{\partial a} = -2 + \frac{2}{q} - \frac{2p_{01}}{p_{11}q}$$

This expression is negative, so $G(a)$ decreases with a.

$$G(0) = c\Big(-1 + \frac{p_{11}}{p_{11}-p_{01}(1-q)} + p_{01}\big(\frac{1}{p_{11}-p_{01}} - \frac{2}{p_{11}q}\big)\Big)$$

We can show that it is negative. So $G(a) < 0$ for any $a < \tilde{a}$.

For low values of the supervision cost ($i.e.$ $a < \tilde{a}$) DS dominates MS.

Second case : for $\bar{a} > a > \tilde{a}$

$G(a) = \frac{p_{11}(c+a)}{p_{11}-p_{01}} + \frac{p_{11}c}{p_{11}-(1-q)p_{01}} - \frac{cp_{11}{p_{01}}^2 q}{(p_{11}-p_{01})^2(p_{11}-(1-q)p_{01})} - 2(c+a) - \frac{2(p_{01}c-(p_{11}-p_{01})a)}{p_{11}q}$

$$\frac{\partial G(a)}{\partial a} = -2 + \frac{p_{11}}{p_{11}-p_{01}} + \frac{2(p_{11}-p_{01})}{p_{11}q} > 0$$

For $\bar{a} = \dfrac{cp_{01}p_{11}q}{(p_{11}-p_{01})(p_{11}-p_{01}(1-q))}$, we have :

$$G(\bar{a}) = \frac{2cp_{01}(1-q)(p_{11}q-(p_{11}-p_{01}))}{p_{11}q(p_{11}-p_{01}(1-q))} > 0$$

For $\tilde{a} = \dfrac{cp_{01}^2 q}{(p_{11}-p_{01})(p_{11}-p_{01}(1-q))}$, we have :

$$G(\tilde{a}) = -\frac{cp_{01}(2p_{01}^2 - 2p_{01}p_{11}(2-q) + 2p_{11}^2 - p_{11}^2 q(2-q))}{(p_{11}-p_{01})p_{11}q(p_{11}-p_{01}(1-q))} < 0$$

So there exists $\bar{\bar{a}} \in [\tilde{a}; \bar{a}]$, such that

- For $a < \bar{\bar{a}}$ DS dominates MS.
- For $a > \bar{\bar{a}}$ MS dominates DS.

Third case : for $\dfrac{cp_{11}}{p_{11}-p_{01}} > a > \bar{a}$

For those values of the monitoring cost, MS dominates DS.

8 Notations

c - cost of effort

a - cost of monitoring

s - the signal observable inside the team

y^H - high team output

y^L - low team output

e_i - agent's i effort

t - side-transfer

M - liability limit

$(1-\lambda)$ - dead weight loss of the side-transfer

w^L - the wage paid to any team-member in case of low output realization

w^H - the wage paid to any team-member in case of high output realization

Δw - bonus ($i.e.$ $\Delta w = w^H - w^L$)

p_{e_i,e_i} - the probability of high output realization conditionally of agents' i and j efforts

r - the probability to observe low signal inside the team when the employee works

q - the probability to observe low signal inside the team when the employee shirks

Bibliographie Générale

Acemoglu D. (1999), "Changes in Unemployment and Wage Inequality : An Alternative Theory and some Evidence", *The American Economic Review*, vol. 89, pp. 1259-1278.

Acemoglu D. and A. Newman (2002), "The Labor Market and Corporate Structure", *European Economic Review*, vol. 46, pp. 1733-1756.

Aghion P. and J. Tirole (1997), "Formal and Real Authority in Organizations", *The Journal of Political Economy*, Vol. 105, pp. 1-29.

Appelbaum E. and R. Batt (1994), *The New American Workplace*, Cornell University Press.

Arnott R. and J. Stiglitz (1991), "Moral Hazard and Non-Market Institutions : Dysfunctional Crowding Out of Peer Monitoring", *American Economic Review*, vol. 81, pp. 179-190.

Autor D., L. Katz and A. Krueger (1998), "Computing Inequality : Have Computers Changed the Labor Market ?", *Quarterly Journal of Economics*, vol. 113, pp. 1169-1213.

Baliga S. and T. Sjostrom (1998), "Decentralization and Collusion", *Journal of Economic Theory*, vol. 83, pp. 196-232.

Boning W., C. Ichniowski and K. Show (2003), "Opportunity Counts : Teams and the Effectiveness of Production Incentives", NBER Working Paper.

Bagnoli M. and T. Bergstrom (2005), "Log Concave Probability and Its Applications", *Economic Theory*, vol. 26, pp. 445-469.

Baker G., R. Gibbons and K.J. Murphy (1994), "Subjective performance measure in Optimal Incentive Contracts", *Quarterly Journal of Economics*, vol. 109, pp. 1125-1156.

Baker G., M. Gibbs and B. Holmström (1994), "The Wage Policy of a Firm", *Quarterly Journal of Economics*, vol. 109, pp. 921-955.

Baumol W., A. Blinder and E. Wolff (2005), "Downsizing in America : Reality, Causes, and Consequences", *Russell Sage Foundation.*

Bewley T. (1999), *Why Wages Don't Fall During a Recession*, Harvard University Press

Blinder A. and D. Choi (1990), "A Shred of Evidence on Theories of Wage Stickiness", *Quarterly Journal of Economics*, vol. 55, pp. 1003-1015.

Bolton P. and M. Dewatripont (1994), "The Firm as a Communication Network", *Quarterly Journal of Economics*, vol. 109, pp. 809-839.

Bresnahan T., E. Brynjolfsson and L. Hitt (2002), "Information Technology, Workplace Organization and the Demand for Skilled Labor : Firm-Level Evidence", *Quarterly Journal of Economics*, vol. 117, pp. 339-376.

Bull C. (1987), "The Existence of Self-Enforcing Implicit Contracts", *Quarterly Journal of Economics*, vol. 102, pp. 147-159.

Cahuc P. and A. Zylberberg (2004), "Labor Economics", *MIT Press.*

Calvo G. and S. Wellisz (1978), "Supervision, Loss of Control and the Optimal Size of the Firm", *Journal of Political Economy*, vol.86, pp. 943-952.

Cappelli P. and K. Chauvin (1991), "An Interplant Test of the Efficiency Wage Hypothesis", *The Quarterly Journal of Economics*, vol. 106, pp. 769-787.

Card D. and A.B. Krueger (1994), "Minimum Wages and Employment : A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania", *The American Economic Review*, vol. 84, pp. 772-793.

Caroli E. and J. Van Reenen (2001), "Skill Biased Organizational Change ? Evidence from a panel of British and French Establishments", *Quarterly Journal of Economics*, vol. 116, pp. 1449-1492.

Chandler A. (1962), "Strategy and Structure", Cambridge, MA : MIT Press

Che Y-K. and S-W. Yoo (2001), "Optimal Incentives in Teams", *American Economic Review*, vol. 81, pp. 525-541.

Coase R. (1937), "The Nature of the Firm", *Economica*, vol. 4, pp. 386-405.

Davis S. and M. Henrekson (1999), "Explaining National Differences in the Size and Industry Distribution of Employment", *Small Business Economics*, vol. 12, pp. 59-83.

Demougin D. and C. Fluet (2001), "Monitoring versus Incentives", *European Economic Review*, vol. 45, pp. 489-496.

Dickens R., S. Machin and A. Manning (1999), "The Effects of Minimum Wages on Employment : Theory and Evidence from Britain", *Journal of Labor Economics*, vol. 17, pp. 1-22.

Garicano L. (2000), "Hierarchies and the Organization of Knowledge in Production", *The Journal of Political Economy*, vol. 108, pp. 874-904.

Garicano L. and E. Rossi-Hansberg (2006), "Organization and Inequality in a Knowledge Economy", *Quarterly Journal of Economics*, vol. 121, pp. 1383-1436.

Gibbons R. (1998), "Incentives in Organizations", *The Journal of Economic*

Perspectives, Vol. 12, pp. 115-132

Gibbons R. and M. Waldman (1999), "Careers in organizations : Theory and evidence", *Handbook of Labor Economics*, in : O. Ashenfelter & D. Card (ed.), vol. 3, pp. 2373-2437

Goldin C. and L. Katz (1999), "The Returns to Skill in the United States Across the Century", *NBER Working Paper* No. 7126.

Hayes, R. and S. Schaefer, (2000), "Implicit Contracts and the Explanatory Power of Top Executive Compensation for Future Performance," *RAND Journal of Economics*, Vol. 31, pp. 273-293.

Holmström B. (1982), "Moral Hazard in Teams", *The Bell Journal of Economics*, vol. 13, pp. 324-340.

Holmström B. and R. Milgrom (1990), "Regulating Trade Among Agents", *Journal of Institutional and Theoretical Economics*, vol. 146, pp. 85-105.

Ichniowski C., G. Prennushi and K. Shaw (1997), "The Effects of Human ressources Management Practices on Productivity", *American economic Review*, vol. 59, pp. 291-313.

Itoh H. (1991), "Incentives to Help in Multi-agents Situations", *Econometrica*, vol. 59, pp. 611-636.

Itoh H. (1993), "Coalitions, Incentives and Risk Sharing", *Journal of Economic Theory*, vol. 60, pp. 410-427.

Kandel E. and E. Lazear (1992), "Peer Pressure and Partnerships", *The Journal of Political Economy*, vol. 100, pp. 801-817.

Katz L. and K. Murphy (1992), "Changes in Relative Wages : Supply and Demand Factors", *Quarterly Journal of Economics*, vol. 107, pp. 35-78.

Kim T. and L.J Taylor (1995), "The Employment Effect in Retail Trade of California's 1988 Minimum Wage Increase", *Journal of Business and Economic Statistics*, vol. 13, pp. 175-182.

Kumar K., R. Rajan and L. Zingales (1999), "What Determines firm size ?", NBER working paper 7208.

Laffont JJ. and P. Rey (2001), "Collusion and Group Lending with Moral Hazard", *IDEI Working Paper.*

Laffont J.J. and J. Tirole (1993), "A Theory of Incentives in Procurement and Regulation", *MIT press*

Lehr B. and F. Lichtenberg (1999), "Information Technology and its Impact on Productivity : Firm-level Evidence from Government and Private Data Sources, 1977-1993", *Canadian Journal of Economics*, vol. 32, pp. 335-362.

Levin J. (2003), "Relational Incentive Contracts", *American Economic Review*, vol. 93, pp. 835-847.

Machin S. and J. van Reenen (1998), "Technology and Changes in Skill Structure : Evidence from Seven OECD Countries", *Quarterly Journal of Economics*, vol. 113, pp. 1215-1244.

Macho-Stadler I. and D. Perez-Castrillo (1993), "Moral Hazard with Several Agents, the Gains from Cooperation", *International Journal of Industrial Organization*, Vol. 11, pp. 73-100.

MacLeod W.B. (2003), "Optimal Contracting with Subjective Evaluation", *American Economic Review*, vol. 93, pp. 216-240.

MacLeod W.B. (2007), "Reputations, Relationships and the Enforcement of Incomplete Contracts" Forthcoming in *Journal of Economic Literature*.

MacLeod W.B. and J.M. Malcomson (1989), "Implicit Contracts, Incentive Compatibility, and Involuntary Unemployment", *Econometrica*, vol. 57, pp. 447-480.

MacLeod W.B. and J.M. Malcomson (1998), "Motivation and Markets", *American Economic Review*, vol. 88, pp. 388-411.

MacLeod W.B. and D. Parent (1997), "Jobs Characteristics and the Form of Compensation", Mimeo, University of Southern California.

Osterman P. (1994), "How Common is Workplace Transformation and Who adopts It ?", *Industrial and Labor Relations Review*, Vol. 47(2), pp. 173-188.

Prendergast C. (1999), "The Provision of Incentives in Firms" *Journal of Economic Literature*, Vol. 37, pp. 7-63.

Radner R. (1992), "Information Processing in Firms and Returns to scale", *Journal of Economic Literature*, vol. 30, pp. 1382-1415.

Radner R. (1993), "The Organization of Decentralized Information processing", *Econometrica*, vol. 61, pp. 1109-1146.

Radner R. and T. Van Zandt (1992), "Information Processing in Firms and Returns to scale", *Annales d'Économie et de Statistique*, vol. 25-26, pp. 265-298.

Rajan R. and J. Wulf (2006), "The Flattening Firm : Evidence from Panel Data on the Changing Nature of Corporate Hierarchies", *The Review of Economics and Statistics*, vol. 88, pp. 759-773.

Roberts J. (2004), "The Modern Firm : Organizational Design for Performance and Growth", *Oxford University Press.*

Shapiro C. and J. Stiglitz (1984), "Equilibrium Unemployment as a Worker Discipline Device", *American Economic Review*, vol. 74, pp. 433-444.

Smith V. (1997), "New Forms of Organization", *Annual Review of Sociology*, Vol. 23, pp. 315-339.

Simon H. (1951), "A Formal Theory of the Employment Relationship", *Econometrica*, vol. 19, pp. 293-305.

Simon H. (1991), "Organizations and Markets", *Journal of Economic Perspectives*, vol. 5, pp. 25-44.

Tirole J. (1992), "Collusion and the Theory of Organizations", *Advances in Economic Theory Sixth World Congress*, vol. 2, pp. 151-207.

Williamson O. (1975), "Markets and Hierarchies : Analysis and Antitrust Implications", New York : Free Press

Résumé

Cette thèse traite trois sujets de théorie des organisations. D'abord, nous étudions l'impact de changements dans l'environnement institutionnel et des caractéristiques du marché du travail sur la structure des organisations, et comment ces évolutions à leur tour affectent l'emploi et les salaires. Ceci est effectué dans le cadre d'un modèle d'équilibre sur le marché du travail, où la structure des organisations est endogène et la production est organisée dans des hiérarchies basées sur les connaissance. Ensuite, nous étudions la forme du contrat incitatif optimal lorsque les employés sont hétérogènes et leur performance n'est pas vérifiable. Nous montrons que l'employeur peut motiver les agents en utilisant simultanément un salaire fixe avec la menace de licenciement en cas d'échec et un bonus basé sur la performance. La part relative de chacun de ces deux outils dans le contrat optimal dépend de l'hétérogénéité des employés, de leur productivité espérée, mais également du taux de rotation exogène et du taux de chômage. Enfin, nous examinons l'impact de la possibilité pour les employés de se superviser mutuellement sur le contrat optimal proposé par l'employeur. Nous montrons que lorsque les employés sont suffisamment bien informés et peu protégés par la responsabilité limitée, cette possibilité de supervision mutuelle permet de réduire le coût des incitations, encouru par l'employeur.

Abstract

This dissertation deals with three topics in the theory of organizations. First, we investigate the impact of institutional and labor market conditions on the structure of organizations and how, in turn, this affects employment and wages. To address these questions we propose an equilibrium model of the labor market, where the structure of organizations is endogenous and production is organized in knowledge based hierarchies. Second, we analyse the optimal incentive contract when employees are heterogenous in their privately observable ability levels, and their performance is observable but not verifiable by third parties. We provide a rationale for the simultaneous use of a termination contract and performance based bonuses. Furthermore, we show that the relative importance of each of these tools in the optimal incentive contract depends on labor market conditions. Third, we investigate the impact of mutual monitoring on the optimal incentive contract proposed by a principal. We show that when agents are sufficiently well informed and protected by a stringent limited liability clause, the possibility for employees to monitor each other reduces the cost of providing incentives.

Discipline : Sciences Économiques (05)

Mots-clés : Incitations, Salaires, Hierarchies, Contrats Relationnels

Laboratoire :

CREST-LEI ; 28, rue des Saints Pères 75007 Paris