

CHAPITRE 2 : INCORPORATION DES CONSIDÉRATIONS DE RÉCIPROCITÉ DANS LA RELATION D'EMPLOI ET NATURE AUTO-ENFORÇANTE DES PRATIQUES DE DON CONTRE DON

Le chapitre précédent a permis de montrer que le processus de décision des agents engagés dans une interaction est motivé par des considérations liées au bien être de l'autre en complément des considérations purement matérielles. Parmi celles-ci, la réciprocité semble jouer un rôle majeur. Les agents punissent des partenaires égoïstes et récompensent ceux qui se sont montrés bienveillants, même si ces deux actions sont coûteuses.

Dans le cas de la relation d'emploi, les considérations de réciprocité représentent des déterminants majeurs des décisions de l'employeur et de l'employé. L'écriture d'un modèle théorique, incorporant directement de telles considérations dans les fonctions d'utilité de ces deux agents, représente un intérêt particulier dans la perspective d'expliquer l'existence de différences de salaire non compensatrices. En effet, il a été montré qu'un comportement réciproque (positif ou négatif) est fondé sur une appréciation préalable non seulement des intentions du partenaire mais aussi de la nature bienveillante de son action. La perception de la bienveillance de l'autre s'appuie sur la comparaison avec une norme considérée comme juste par le joueur. Dès lors que les comportements sont basés sur une telle comparaison, il faut s'attendre à ce que la norme en question exerce une attraction particulière sur les choix des partenaires. Les politiques de rémunération devraient donc respecter une certaine compatibilité vis à vis de cette norme pour assurer la réalisation du niveau d'effort désiré par les salariés (celui-ci pouvant être également considéré comme la norme d'effort légitimement revendiquée par l'employeur). Or, ces deux normes peuvent différer d'un établissement à l'autre, suivant l'historique des relations entre employeur et employé, le secteur impliqué ou la

nature des tâches demandées à l'employé. Ainsi, la nature spécifique de ces normes et l'importance des considérations de réciprocité justifieraient la mise en place de pratiques salariales individualisées et donc l'observation de différences de salaire non compensatrices au sein de la firme.

Le modèle théorique proposé précise les conditions de mise en œuvre, en termes de motivation pour la réciprocité, de pratiques d'échange de dons stables entre employeur et employé. L'hétérogénéité des motivations intrinsèques des agents est prise en compte. Les équilibres théoriques de l'interaction sont définis selon la qualité de l'appariement entre employeur et employé ainsi que la nature des arbitrages effectués par ces derniers entre recherche de gain et comportements réciproques et suivant les caractéristiques de la relation d'emploi.

Une première section est consacrée à l'exposé de diverses fonctions d'utilité modifiées présentes dans la littérature et susceptibles d'être pertinentes pour la modélisation de la relation d'emploi. En accord avec les hypothèses 1 et 2 du premier chapitre, la modification doit permettre de tenir compte des intentions des joueurs (et des croyances sur les intentions), de leur hétérogénéité en termes de considérations intrinsèques de réciprocité, et, enfin de l'arbitrage entre motivations intrinsèques et extrinsèques. Dans la section 2, un jeu séquentiel dans lequel l'employeur prend sa décision en premier est défini. Une telle structure du jeu s'impose, car dans la relation d'emploi, l'employé connaît les conditions salariales de son contrat avant d'exercer ses propres obligations. Le jeu d'échange de cadeaux séquentiel est paramétré en vue de tenir compte des particularités de la relation d'emploi. Les gains des joueurs ne sont ni donnés, ni exogènes, mais représentés par des fonctions. Pour l'employeur, la fonction de gain matériel se définit par le surplus entre la profitabilité de l'employé et le salaire offert, pondéré par le niveau d'effort mis en œuvre. Pour l'employé, cette fonction se résume à la différence entre le salaire et le coût lié à l'effort choisi.

La modélisation de Dufwenberg & Kirchsteiger (1998) permet alors de décrire l'utilité des joueurs à chaque nœud du jeu dans l'espace des stratégies disponibles (section 3).

Une telle démarche permet *in fine* (section 4) de décrire les équilibres théoriques possibles du jeu en fonction du degré de motivation intrinsèque des joueurs pour la réciprocité. Plus largement, elle permet de prédire les résultats à attendre de la confrontation entre chaque type de joueurs, plus ou moins égoïstes, compte tenu des caractéristiques de la relation d'emploi envisagée. La présence d'un agent égoïste induit-elle obligatoirement un échec de l'interaction ? Suffit-il d'être « mutuellement réciproques » pour se comporter

systématiquement de façon bienveillante ? Quelle est la robustesse des motivations sociales vis à vis de variations des paramètres caractérisant la relation d'emploi ? La réponse à l'ensemble de ces questions détermine le crédit à accorder à l'explication des différences non compensatrices de salaire proposée par Akerlof et Yellen (1990) tout en permettant d'évaluer l'efficacité à attendre de pratiques généralisées de don contre don dans la relation d'emploi.

1. MODIFICATION DES FONCTIONS D'UTILITÉ ET IMPORTANCE DES INTENTIONS DANS LA MODÉLISATION DE LA RELATION D'EMPLOI

1.1. Débat sur l'opportunité de modifier les fonctions d'utilité

Ce chapitre permet d'apporter un fondement théorique à l'observation d'un fait stylisé précis (l'existence de différences non compensatrices de salaire) mais également de répondre à la question de l'efficacité de la mise en place de pratiques de don contre don dans la détermination des salaires. Pour ce faire une approche pragmatique est privilégiée. Cette volonté conduit à proposer une modélisation de la relation d'emploi fondée sur une modification à la marge des fonctions d'utilité de l'employeur et de l'employé permettant d'incorporer la composante sociale de leurs motivations. Ce faisant, nous sommes amenés à prendre partie dans le débat qui oppose les expérimentalistes sur l'intérêt d'apporter de telles modifications. En effet, dès lors que les comportements observés dans l'ultimatum ne peuvent être expliqués par de simples pratiques de maximisation des gains matériels, Ochs & Roth (1989) ou Bolton (1991) suggèrent d'incorporer un argument relatif dans les fonctions d'utilité pour tenir compte de l'importance accordée par les joueurs aux gains de leur partenaire. Une telle démarche a l'avantage indéniable de permettre l'utilisation des outils traditionnels d'optimisation à des fins prédictives.

Cependant, certains auteurs comme Güth (1995) dénoncent l'aspect artificiel d'une telle démarche qui s'apparente plutôt à un « banc de réparation » de la théorie des jeux :

"Very often, this type of research resembles, however, a neoclassical repair shop in the sense that one first observes behaviors for a certain environment and then defines a suitable optimization or game model which can account for what has been observed"⁷⁴

Citation 2-1: Güth, 1995, p 342.

Ainsi, l'adjonction de nouveaux arguments dans la fonction d'utilité ne produirait pas d'explication satisfaisante des comportements des agents. Selon Güth, leur comportement est plutôt le fruit d'un processus dynamique de prise de décision dans un contexte de rationalité limitée. Dans une première étape, le joueur tire partie de son expérience passée pour analyser la situation dans laquelle il se trouve. Le but est de répéter un comportement qui s'est déjà avéré fructueux en pareille situation. Cette étape correspond à une exigence minimale de rationalité puisqu'elle consiste simplement à mettre en relation histoire du joueur et situation présente. Une deuxième étape représente la mise en perspective stratégique du joueur par rapport à l'autre. S'il s'agit de l'offreur, cette étape consiste à déterminer quelle serait la demande maximale qu'il pourrait faire sans avoir à supporter un refus de la part du répondant. Une fois déterminé le seuil d'acceptation du répondant, le joueur soumet ce partage à un test d'acceptabilité : le partage est-il socialement acceptable ? Est-il en accord avec les normes de comportement du joueur ou du groupe ? Va-t-il entraîner une réaction négative de la part du répondant ? Enfin une troisième étape, représentant un degré supérieur de rationalité, correspond à un processus d'induction à rebours de la part de l'offreur pour évaluer la réelle disposition du répondant, guidé par ses émotions, à refuser le partage qui lui est proposé.

Si une telle interprétation constitue une voie intéressante pour l'explication des comportements des sujets, elle pose néanmoins des difficultés analytiques évidentes quant à la capacité de prédiction des équilibres du jeu ou du comportement moyen des sujets. Chaque joueur diffère dans ses propres expériences, dans sa capacité à mobiliser la technique de l'induction à rebours., si bien qu'en l'état actuel de la théorie behavioriste proposée par Güth, il paraît difficile de dégager pour un jeu donné, une tendance générale des comportements.

Il apparaît être plus intéressant d'adopter le corpus qui consiste à ajouter des arguments supplémentaires à la fonction d'utilité. Cependant, il importe de garder à l'esprit les limites potentielles d'une telle démarche ainsi que le risque non négligeable de « bricolage » ad hoc de la théorie néoclassique. L'utilisation d'une telle démarche se doit donc d'être prudente.

⁷⁴ « Très souvent, on peut assimiler ce genre de recherches à une sorte d'atelier de réparation de la théorie néoclassique dans le sens où l'on observe d'abord les comportements dans un certain environnement puis on définit le type d'optimisation ou le modèle le plus adapté pour rendre compte de ce qui a été observé »

C'est pourquoi, un soin particulier a été apporté à justifier la nature des arguments introduits dans la modélisation par le biais d'observations robustes et stables effectuées en environnement contrôlé.

1.2. Le choix de la fonction d'utilité modifiée pertinente pour la modélisation de la relation d'emploi

De fait, les considérations de réciprocité peuvent, pour partie, rendre compte des déviations systématiquement observées par rapport à l'équilibre théorique des jeux de négociation bilatérale. Suivant la voie initiée par Ochs & Roth (1989) et Bolton (1991), un certain nombre de travaux permettent d'incorporer directement ce type de préoccupations dans les fonctions d'utilité. En schématisant, il est possible de les regrouper en deux grandes catégories. La première est constituée des fonctions d'utilité dont la partie reflétant les préoccupations sociales des joueurs est essentiellement axée sur des considérations distributives de fin de jeu. C'est le cas des fonctions développées par Bolton & Ockenfels (1997)⁷⁵ et Fehr & Schmidt (1997)⁷⁶. La deuxième catégorie correspond aux fonctions d'utilité induisant une réaction possible du joueur par rapport aux intentions qu'il attribue à son partenaire lorsqu'il prend une décision. Ici, ce n'est pas tant la distribution finale que la variété des choix disponibles avant la prise de décision qui influence la réaction positive ou négative du joueur. Le jeu devient un jeu psychologique dans lequel les croyances jouent un rôle prépondérant. Cette voie est initiée par Geanakoplos, Pearce & Stacchetti (1989)⁷⁷ et a été développée par Rabin (1993)⁷⁸ et Dufwenberg & Kirchsteiger (1998)⁷⁹. Le paragraphe suivant montre la pertinence de

⁷⁵ Bolton G.E., Ockenfels A., (1997), « ERC, A theory of Equity, Reciprocity and Competition », *Working Paper, August.*; *American Economic Review* (2000), 90, 1, March, pp. 166-193.

⁷⁶ Fehr E., Schmidt K.M., (1997), « A theory of fairness, competition and cooperation », *Working Paper, December.*

⁷⁷ Geanakoplos J., Pearce D., Stacchetti E., (1989), « Psychological games and sequential rationality », *Games and Economic Behavior*, 1, pp 60-79.

⁷⁸ Rabin M., (1993), « Incorporating fairness into game theory and economics », *American Economic Review*, vol 83 n°5, pp 1281-1302.

⁷⁹ Dufwenberg M., Kirchsteiger G., (1998), « A theory of sequential reciprocity », *Working Paper, Economic Science Association annual meeting, Mannheim, Germany.*

l'utilisation d'une fonction d'utilité appartenant à cette dernière catégorie pour la modélisation de la relation d'emploi.

1.2.1. Eléments de choix de la fonction d'utilité modifiée

Le tableau suivant résume les deux catégories de fonctions d'utilité modifiées utilisées dans la littérature pour traiter du cas de la détermination des salaires. Les fonctions de Bolton & Ockenfels (1997) et Fehr & Schmidt (1997) sont néanmoins présentées bien qu'elles ne tiennent pas compte explicitement des considérations de réciprocité mais plutôt de l'aversion auto centrée pour l'inégalité (*self centered inequality aversion*). Cependant, leurs auteurs se montrent particulièrement revendicatifs quant au pouvoir explicatif de ce type de considération, même pour les jeux dans lesquels la réciprocité semble fournir l'explication la plus plausible (jeux d'investissement, échange de cadeaux...). Le choix de la fonction d'utilité modifiée adaptée à la modélisation de la relation employeur-employé passe par une remise en cause préalable de cette première catégorie de fonction d'utilité. L'argument principal est que ces fonctions négligent d'incorporer un élément prépondérant dans l'explication des comportements de négociation de l'employeur et de l'employé, à savoir l'influence des intentions.

Tableau 2-1 : Fonctions d'utilité modifiées disponibles dans la littérature.

Auteurs	Forme générale de la fonction	Fonction dans un jeu à deux joueurs i et j
E.R.C. Bolton & Ockenfels (1997)	$U_i = v_i(\pi_i, \lambda_i)$ $\lambda_i = \frac{n\pi_i}{c}$	$U_i = v_i(\pi_i, \lambda_i)$ $\lambda_i = \frac{2\pi_i}{c}$ <p>expl : $u_i(\pi_i, \lambda_i) = a_i\pi_i - \frac{b_i}{2}(\lambda_i - \frac{1}{2})^2$</p>
F.C.C. Fehr & Schmidt (1997)	$U_i = x_i - \alpha_i \frac{1}{n-1} \sum_{j \neq i} \max\{\pi_j - \pi_i, 0\}$ $- \beta_i \frac{1}{n-1} \sum_{j \neq i} \max\{\pi_i - \pi_j, 0\}$	$U_i = x_i - \alpha_i \max\{\pi_j - \pi_i, 0\}$ $- \beta_i \max\{\pi_i - \pi_j, 0\}$
Levine (1996)	$U_i = \pi_i + \sum_{j \neq i} \frac{a_i + \lambda a_j}{1 + \lambda} \pi_j$ $0 \leq \lambda \leq 1$	$U_i = \pi_i + \frac{a_i + \lambda a_j}{1 + \lambda} \pi_j$ $0 \leq \lambda \leq 1$
Reciprocal Fairness Rabin (1993)	$U_i = \pi_i(a_i, b_j) + \tilde{f}_j(b_i, c_j) [1 + f_i(a_i, b_j)]$ <p>(défini dans un environnement à deux joueurs)</p> <p>avec,</p> $\tilde{f}_j(\cdot) = \frac{\pi_i(c_j, b_i) - \pi_i^e(c_j)}{\pi_i^h(c_j) - \pi_i^{\min}(c_j)}$ $f_i(\cdot) = \frac{\pi_j(b_j, a_i) - \pi_j^e(b_j)}{\pi_j^h(b_j) - \pi_j^{\min}(b_j)}$	
Falk & Fischbacher (1998)	$U_i(f) = \pi_i(f) + \rho_i \sum_{n \rightarrow f} \varphi(n) \sigma(n, f)$ <p>(défini dans un environnement à deux joueurs)</p> <p>avec,</p> <p>f : nœud terminal du jeu</p> <p>$\sigma(n, f)$: terme de réciprocité (réaction)</p> <p>$\varphi(n) = \mathcal{G}(n) \Delta(n)$, où $\Delta(n)$ représente la comparaison des payoffs et $\mathcal{G}(n)$ une pondération dont la valeur est déterminée par l'intention attribuée à l'autre.</p>	
S.R.E. Dufwenberg & Kirchsteiger (1998)	$U_i(\cdot) = \pi_i(a_i, (b_{ij})_{j \neq i})$ $+ Y_i \sum_{j \in N \setminus \{i\}} \kappa_{ij}(a_i, b_{ij}) \lambda_{ij}(b_{ij}, (c_{ijk})_{j \neq k})$	$U_i(\cdot) = \pi_i(a_i, (b_{ij})_{j \neq i})$ $+ Y_i \kappa_{ij}(a_i, b_{ij}) \lambda_{ij}(b_{ij}, c_{iji})$ <p>avec,</p> $\kappa_{ij}(\cdot) = \pi_j(a_i, b_{ij}) - \pi_j^{e_i}(b_{ij})$ $\lambda_{ij}(\cdot) = \pi_i(b_{ij}, c_{iji}) - \pi_i^{e_j}(c_{ijk})$

Les fonctions de Fehr & Schmidt (FCC) et Bolton & Ockenfels (ERC), comme l'ensemble des fonctions d'utilité modifiées, incorporent le gain matériel du joueur sous une forme additive. Sa motivation sociale est représentée par une comparaison entre la part relative du surplus qu'il reçoit et la part relative qui devrait lui être attribuée compte tenu du nombre de joueurs et de la taille du gâteau. La base de la comparaison se fait sur le principe d'équité. La satisfaction du joueur est donc dégradée s'il constate une différence à son désavantage par

rapport à une telle part relative. Dans *FCC*, cette préoccupation est représentée par le terme

$$\alpha_i \frac{1}{n-1} \sum_{j \neq i} \max\{\pi_j - \pi_i; 0\}$$
 alors que λ_i remplit cette fonction dans *ERC*.

Pourquoi parler d'aversion auto centrée pour l'inégalité ? Compte tenu de la modélisation adoptée par Bolton & Ockenfels (1997), on constate que l'utilité du joueur n'est pas altérée s'il reçoit une part plus importante du surplus. Dans *FCC*, un troisième terme tient compte de cette diminution de la satisfaction du joueur s'il se trouve trop bien traité par rapport aux autres. Cependant l'hypothèse $\alpha_i > \beta_i$ impose qu'à écart donné, la perte de satisfaction totale soit supérieure si la différence est à son détriment plutôt qu'à son avantage.

De plus, dans ces deux fonctions l'inégalité des gains entre les autres est supposée n'avoir aucune conséquence comportementale sur le joueur *i* considéré.

Dans ces formulations (ainsi que dans celle de Levine, 1996), l'accent est mis sur la distribution finale des gains entre les participants au jeu. Ainsi, à part relative donnée, le comportement d'un joueur reste le même, que cette part provienne d'une action volontaire du partenaire ou d'un processus aléatoire. Or l'évidence expérimentale aussi bien qu'un certain nombre de faits stylisés de la relation d'emploi montrent que les joueurs tendent à être plus indulgents vis à vis de leur partenaire s'ils sont incapables de lui attribuer une quelconque responsabilité dans la survenance d'une issue défavorable. Concernant l'occurrence d'une action bienveillante, au contraire les joueurs réagissent d'autant plus positivement qu'ils sont sûrs que leur situation provient d'une action intentionnelle de la part de leur partenaire (Cf. Chapitre précédent, section 2).

Dès lors que de tels éléments liés à l'intentionnalité des actions jouent un rôle comportemental important, *ERC* et *FCC* ne peuvent constituer une formulation pertinente pour la modélisation du jeu d'échange de cadeaux dans la relation d'emploi. Par ailleurs, l'hypothèse d'indifférence comportementale du joueur *i* vis à vis de la distribution du gâteau entre les autres joueurs est remise en cause par les résultats expérimentaux de Bolle & Kritikos (1998)⁸⁰.

Les formulations suivantes, initiés par les travaux de Geanakoplos, Pearce & Stacchetti (1989), permettent directement d'incorporer les éléments liés à l'intentionnalité. Il reste donc à

⁸⁰ Bolle F., Kritikos A., (1998), « Self-Centered Inequality Aversion versus Reciprocity and Altruism », papier présenté à la conférence annuelle de l'ESA, Mannheim, juin.

définir la formulation la plus pertinente pour modéliser les comportements de l'employeur et de l'employé dans le contexte de la détermination des salaires.

1.2.2. Jeu psychologique et importance des intentions

En considérant que l'action des joueurs est conditionnée pour partie par une réaction émotive vis à vis du choix de leur partenaire (Cf. Blount (1995), Ruffle, 1998⁸¹) : récompenser une action bienveillante et sanctionner une action malveillante, il importe de préciser la manière dont ces joueurs jugent l'autre. Par conséquent, l'analyse doit porter non seulement sur l'aspect distributif mais aussi sur les intentions qui ont déterminé les choix :

"A crucial feature of our theory is the way we capture how people evaluate whether an experienced action is kind or unkind. Two aspects are essential here, namely (i) the outcome or the consequences of an action and (ii) the underlying motivation, i.e., the intentions involved"⁸²

Citation 2-II: Falk A., Fischbacher U., 1998, p 3⁸³.

L'idée de base de Geanakoplos, Pearce & Stacchetti (1989) consiste à construire un jeu psychologique à partir du jeu de référence résumé par les gains matériels associés à chaque nœud terminal du jeu. Les gains des joueurs sont alors modifiés pour tenir compte non seulement de leurs actions mais aussi de leurs croyances. Ils définissent alors un *équilibre de Nash psychologique*, dont les propriétés sont similaires à l'équilibre de Nash à ceci près que pour satisfaire aux conditions d'existence, les croyances des joueurs doivent être en accord avec le comportement sélectionné.

Rabin (1993) propose une formulation de la fonction d'utilité des joueurs incorporant de tels éléments psychologiques dans le cadre des jeux simultanés (voir Tableau 2-1 pour la forme de la fonction). L'utilité attendue par le joueur dépend alors de trois éléments : son ensemble de stratégies, ses croyances sur l'ensemble des stratégies de son partenaire et enfin ses croyances sur la croyance de l'autre sur ses propres stratégies. La fonction d'utilité est alors composée

⁸¹ Ruffle B.j., (1998), « More is better, but fair is fair : tipping in dictator and ultimatum games », *Games and Economic Behavior*, vol 23, pp 247-265.

⁸² « Un élément crucial de notre théorie repose sur la modélisation de la façon dont les agents interprètent une action comme bienveillante ou non. Deux éléments sont déterminants, (i) le résultat ou les conséquences d'une action et (ii) les motivations, i.e. les intentions qui les sous tendent »

⁸³ Falk A., Fishbacher U., (1998), « A theory of reciprocity », *Working Paper*, University of Zurich, November.

par les gains matériels du jeu auxquels s'ajoutent des gains psychologiques dépendant de deux facteurs : la perception de la bienveillance de l'autre et, en retour, l'expression d'une action bienveillante ou malveillante dans l'espace des stratégies disponibles et efficaces du joueur. La bienveillance d'un joueur i vis à vis de j est donnée par la fonction

$$f_i(.) = \frac{\pi_j(b_j, a_i) - \pi_j^e(b_j)}{\pi_j^h(b_j) - \pi_j^{\min}(b_j)}. \text{ Elle consiste en une comparaison entre le gain attendu du jeu}$$

(sachant que j est supposé choisir la stratégie b_j) et une norme de référence considérée comme juste par le joueur. Cette distribution de référence est déterminée ici sur le principe d'équité, elle correspond à la moyenne arithmétique entre le plus élevé et le plus faible des gains Pareto-éfficiants susceptibles d'être attribués à j par l'action de i :

$$\pi_j^e(b_j) = \frac{1}{2} [\pi_j^h(b_j) + \pi_j^\lambda(b_j)]. \text{ Cette comparaison est pondérée par la différence entre le gain}$$

potentiel le plus élevé de j et son gain le plus faible du jeu. La bienveillance du joueur i envers j est donc nulle si j obtient le partage équitable. Si i revendique une part plus importante, $f_i(.)$ devient négative, reflétant une action malveillante vis à vis de j .

La perception de la bienveillance de j envers i est calculée de la même façon et est représentée

$$\text{par } \tilde{f}_j(.) = \frac{\pi_i(c_j, b_i) - \pi_i^e(c_j)}{\pi_i^h(c_j) - \pi_i^{\min}(c_j)}. \text{ Dès lors que l'on normalise les deux fonctions de}$$

bienveillance, leurs valeurs sont comprises dans l'intervalle $\left[-1; \frac{1}{2}\right]$. Le joueur i maximise

donc une fonction d'utilité définie par :

$$U_i = \pi_i(a_i, b_j) + \tilde{f}_j(b_i, c_j) [1 + f_i(a_i, b_j)],$$

où les fonctions de bienveillance apparaissent comme indépendantes des transformations affines de l'échelle des gains matériels.

Considérant une telle fonction, la meilleure réponse du joueur i à une action anticipée comme malveillante (bienveillante) est représentée par la punition (récompense), même si ces pratiques sont coûteuses pour le joueur. Une telle fonction d'utilité offre l'avantage de rendre rationnelles et crédibles de telles pratiques coûteuses mais comporte un certain nombre de limites pour la modélisation du jeu d'échange de cadeaux entre employeur et salarié. Compte tenu de la forme des gains psychologiques, une augmentation du rang des gains matériels diminue mécaniquement l'influence de ces considérations émotives. De plus, cette fonction d'utilité est définie pour des jeux non séquentiels contrairement au jeu d'échange de cadeaux

entre employeur et employé. En effet, lors de l'établissement du contrat, l'employé connaît sa rémunération. L'incomplétude porte sur le niveau d'effort sélectionné par l'employé.

La modélisation de Dufwenberg & Kirchsteiger (1998) permet de dépasser les limitations du concept de Rabin (1993) tant en ce qui concerne le rôle de la partie psychologique des gains lorsque les enjeux augmentent que sur son applicabilité aux jeux séquentiels. La transposition du concept de Rabin aux jeux sous forme extensive implique que les joueurs sont amenés à réviser leurs croyances suivant les nœuds du jeu dans lesquels ils se trouvent. Dès lors que dans un jeu psychologique, le comportement plus ou moins bienveillant d'un joueur i dépend de son anticipation sur les intentions de l'autre et que ces dernières dépendent elles-mêmes de croyances, sa motivation pour la réciprocité dépend de croyances sur des croyances (croyance de second ordre). La nature de ces croyances dépend donc fortement du sous-jeu dans lequel les joueurs se trouvent au moment de leur choix. Il est par conséquent nécessaire d'opérer une distinction entre les croyances initiales des joueurs et les révisions qu'il sont amenés à faire tout au long du jeu. Cette distinction constitue la différence principale par rapport aux jeux psychologiques de Geanakoplos & al. (1989).

Si $N = \{1, 2, \dots, n\}$ représente le nombre de joueurs tel que $n \geq 2$, on désigne A_i comme l'ensemble des stratégies à la disposition du joueur i . On définit $\pi(\cdot)$, comme la fonction de gains matériels du joueur i associée à chaque ensemble de stratégie au nœud terminal du jeu : $\pi_i : A \rightarrow R$. Des gains psychologiques tenant compte du soucis pour la réciprocité du joueur i sont ajoutés à cette fonction de gains matériels. Comme dans Rabin (1993), c'est à ce niveau que sont introduites les croyances des joueurs. On appelle $B_{ij} = A_j$, l'ensemble des croyances possibles du joueur i sur les stratégies du joueur j avec lequel il est apparié. On définit également $C_{ijk} = B_{jk} = A_k$ comme l'ensemble des croyances de i sur la croyance de j sur la stratégie d'un joueur k (dans un contexte à deux joueurs, C_{iji} s'écrit comme l'ensemble des croyances de i sur les croyances de j sur les stratégies de i). En termes de gains, si i choisit une action $a_i \in A_i$ croyant que les autres joueurs se comportent selon le profil de stratégies décrit par $(b_{ij})_{j \neq i} \in B_{ij}$, alors i pense procurer un gain matériel $\pi_j(a_i, b_{ij})$ à j compte tenu du nœud r dans lequel il se trouve au moment de sa décision. Compte tenu de la nature séquentielle du jeu, le gain que i attribue à j par son action fait partie de l'ensemble des gains de fin de jeu qui restent possibles à ce moment du jeu : $\{\pi_j(a'_i, b_{ij}) / a'_i \in A_i\}$. A cet égard, la base de comparaison pour déterminer si i a été bienveillant ou non dans son choix de comportement est représentée par cet ensemble de gains possibles et non plus par l'ensemble

des gains matériels de fin de jeu. Il y a identité entre ces deux ensembles uniquement lorsque les joueurs se situent au premier nœud de décision du jeu. Cette particularité est la conséquence directe de la possibilité donnée aux joueurs de réviser leurs croyances dans les jeux en forme extensive.

Les gains psychologiques comportent trois éléments. Un premier élément représente la préférence intrinsèque du joueur i pour la réciprocité et est noté Y_i . Elle revient à pondérer les gains extra matériels. Si $Y_i=0$, le joueur n'exhibe aucune motivation pour la réciprocité et se comporte de manière à maximiser son gain matériel de fin de jeu. Les deux autres éléments sont constitués d'une partie *anticipation/observation* de l'action de l'autre, visant à déterminer la bienveillance de l'action de ce dernier, et d'une partie *action/réaction*, qui évalue la bienveillance de l'action de l'autre vis à vis du joueur i .

Afin de définir la bienveillance de i par rapport à j (*action/réaction*), on utilise la même référence que Rabin (1993) par rapport au gain équitable, qui représente la norme psychologique des joueurs. La fonction de bienveillance prend la valeur 0 si i procure à j le gain équitable dans l'ensemble des gains accessibles au nœud de décision considéré. Ce gain est noté $\pi_j^{e_i}$ et correspond à la moyenne arithmétique entre le gain le plus élevé et le plus faible de cet ensemble⁸⁴. La base de comparaison est l'équité. $\pi_j^{e_i}$ est défini de la façon suivante :

$$\pi_j^{e_i} = \frac{1}{2} \left[\max \{ \pi_j(a_i, b_{ij}) / a_i \in E_i \} + \min \{ \pi_j(a_i, b_{ij}) / a_i \in E_i \} \right],$$
 où E_i est l'ensemble des stratégies efficientes.

La fonction de bienveillance de i par rapport à j s'écrit alors :

Fonction de bienveillance de i envers j :

$\kappa_{ij}(a_i, b_{ij}) : A_i \times B_{ij} \rightarrow R$ telle que

$$\kappa_{ij}(a_i, b_{ij}) = \pi_j(a_i, b_{ij}) - \pi_j^{e_i}(b_{ij})$$

On procède de la même façon pour la partie *anticipation/observation* du joueur i qui conditionne son comportement. Il doit donc former une anticipation sur la bienveillance de

⁸⁴ On suppose que les joueurs choisissent des stratégies dans l'ensemble des stratégies efficaces. Une stratégie est considérée comme inefficace s'il existe une autre stratégie qui, compte tenu de l'histoire du jeu et des décisions qui ont déjà été prises, ne procure pas de gains inférieurs pour tous les joueurs et permet à au moins un des joueurs d'améliorer potentiellement sa situation matérielle.

l'action de son partenaire. Cette anticipation dépendant des intentions qu'il lui attribue. λ_{iji} se définit alors comme la différence entre le gain que le joueur **i** pense que **j** veut lui attribuer et le gain équitable qu'il est en droit d'attendre compte tenu du nœud de décision atteint.

Croyance de **i** sur la bienveillance de **j** :

$\lambda_{iji}(b_{ij}, c_{ijk}) : B_{ij} \times C_{ijk} \rightarrow R$ telle que

$$\lambda_{iji}(b_{ij}, c_{iji}) = \pi_i(b_{ij}, c_{iji}) - \pi_i^{e_j}(c_{ijk})$$

Compte tenu des éléments matériels et psychologiques, la fonction d'utilité générale du joueur **i** s'écrit :

Fonction d'utilité du joueur **i** :

$U_i : A_i \times (B_{ij} \times C_{ijk}) \rightarrow R$ telle que

$$U_i(a_i, b_{ij}, c_{ijk}) = \pi_i(a_i, (b_{ij})_{j \neq i}) + Y_i \sum_{j \in N / \{i\}} \kappa_{ij}(a_i, b_{ij}) \lambda_{iji}(b_{ij}, (c_{ijk})_{j \neq k})$$

si le jeu comporte deux joueurs :

$$U_i(a_i, b_{ij}, c_{iji}) = \pi_i(a_i, (b_{ij})_{j \neq i}) + Y_i \kappa_{ij}(a_i, b_{ij}) \lambda_{iji}(b_{ij}, c_{iji})$$

Si **i** anticipe une action malveillante de **j**, il s'attend à un gain inférieur à la norme et par conséquent λ_{iji} est négatif. Sa réaction consiste à adopter lui même un comportement malveillant rendant κ_{ij} également négatif. Les trois éléments des gains psychologiques étant introduits multiplicativement, le joueur **i** augmente sa satisfaction à se comporter ainsi dans une proportion déterminée par sa préférence individuelle pour la réciprocité (Y_i). A l'inverse, un joueur motivé par la réciprocité augmente sa satisfaction en récompensant une action bienveillante de la part de son partenaire. La modification des fonctions d'utilité afin de prendre en compte les motivations sociales des joueurs permet de rendre rationnel tout comportement de punition ou de récompense. D'une part, il importe de transposer ce type de modélisation à la relation d'emploi en permettant aussi bien à l'employeur qu'à l'employé d'être motivés par des considérations de réciprocité. Et d'autre part, il est nécessaire d'analyser dans quelle mesure cette motivation peut représenter un mécanisme d'implémentation de la coopération.

2. DÉFINITION D'UN JEU D'ÉCHANGE DE DONNS ENTRE EMPLOYEUR ET EMPLOYÉ

Le modèle utilisé correspond à une version basique du jeu d'échange de cadeaux à laquelle une paramétrisation a été introduite pour tenir compte des caractéristiques de la relation d'emploi. L'équilibre de Nash d'un tel jeu correspond au résultat le plus défavorable collectivement pour les deux protagonistes à la négociation. Il correspond à un effort minimum de la part de l'employé et à une offre de salaire minimale proposée par l'employeur. L'utilisation d'une telle structure de jeu séquentiel à un coup ne laisse aucune ambiguïté sur l'issue du jeu impliquant des individus "égoïstes". L'interaction n'étant pas répétée, aucune possibilité n'est donnée aux joueurs d'investir en réputation et, par conséquent, de coopérer pour des raisons de stratégie de long terme. Cette démarche permet d'assurer que les comportements coopératifs observés seront la conséquence directe de l'introduction des considérations de réciprocité dans les choix des joueurs.

2.1. Ecriture du jeu d'échange de cadeaux sous forme extensive : hypothèses

Dans le jeu d'échange de cadeaux, l'employeur peut proposer un salaire élevé \bar{w} , correspondant au don, ou un salaire faible \underline{w} , reflétant une rémunération à la productivité marginale qui lui laisse la totalité du surplus de l'interaction. Il prend sa décision compte tenu de la profitabilité q de l'employé qu'il connaît. Ce paramètre représente le gain potentiel maximal que l'employeur peut retirer de son salarié compte tenu de ses caractéristiques individuelles, du type de tâche effectuée. Il est considéré comme exogène. Face à une telle proposition de salaire, l'employé choisit alors son effort. Par soucis de simplification, l'effort est supposé prendre deux valeurs possibles : une valeur faible \underline{e} qui correspond à un effort nul ($\underline{e} = 0$) de la part du salarié, associée à un coût de l'effort nul $C(\underline{e}) = 0$; et une valeur élevée \bar{e} correspondant à un effort maximum lui imposant un coût $C(\bar{e})$.

Le gain matériel du salarié maximisant son effort correspond donc à $\bar{w} - C(\bar{e})$ s'il reçoit une offre de salaire élevée et à $\underline{w} - C(\bar{e})$ s'il reçoit une proposition faible. De la même façon, son gain devient respectivement \bar{w} ou \underline{w} s'il dégrade son effort.

Comme le jeu est décrit en termes de surplus, l'employeur reçoit quant à lui la différence entre la profitabilité de l'employé et le salaire qu'il choisit de lui verser, le tout pondéré par l'effort effectivement mis en œuvre par l'employé dans la relation d'emploi⁸⁵.

Le jeu sous forme extensive peut donc s'écrire ainsi :

Graphe 2-1 : Jeu d'échange de cadeau entre employeur et employé.

$$\begin{array}{cccc} \pi^E = 0 & \pi^E = \bar{e}(q - \underline{w}) & \pi^E = 0 & \pi^E = \bar{e}(q - \bar{w}) \\ \pi^L = \underline{w} & \pi^L = \underline{w} - C(\bar{e}) & \pi^L = \bar{w} & \pi^L = \bar{w} - C(\bar{e}) \end{array}$$

Où π^E et π^L représentent respectivement les gains de l'employeur et de l'employé pour chacune des éventualités liées à l'interaction.

Dans ce jeu, l'employeur joue en premier par soucis de réalisme vis à vis des pratiques adoptées couramment sur le marché du travail. En effet, l'employé connaît la rémunération qu'il percevra au moment où il choisit son niveau d'effort puisque cet élément est inclus dans le contrat. Par contre, l'inobservabilité de l'effort rend le contrat incomplet de sorte que tout niveau minimum requis par l'employeur ne représente qu'un *signal gratuit*⁸⁶ du point de vue du salarié.

L'introduction d'une telle séquentialité dans les prises de décision de l'employeur et de l'employé rend le jeu différent de celui de Akerlof (1982)⁸⁷. En effet, chez Akerlof (1982), les deux joueurs prennent leur décision simultanément, le jeu correspond ainsi à un jeu classique de dilemme du prisonnier.

⁸⁵ Nous écrivons le gain de l'employeur ainsi afin de restreindre l'ensemble de ses gains dans \mathfrak{R}^+ . Si l'on ne pondérait que la profitabilité de l'employé (q) par le niveau d'effort, ses gains pourraient être négatifs si l'employé ne coopère pas. Dès lors que l'on modifie les fonctions d'utilité, il faudrait alors tenir compte du phénomène d'aversion aux pertes dans la modélisation (*loss aversion*), ce qui compliquerait inutilement le modèle.

⁸⁶ Traduction de *cheap talk*.

⁸⁷ Akerlof G., (1982), « Labor contracts as partial gift exchange », *Quarterly Journal of Economics*, 97, pp 543-569

2.2. Equilibre de Nash du jeu impliquant des joueurs égoïstes

Compte tenu d'une telle structure, et puisque les joueurs sont supposés chercher uniquement à maximiser leurs gains de fin de jeu, l'équilibre de Nash parfait en sous-jeu est représenté par un échec de l'échange de cadeau qui conduit à une proposition de salaire faible associée à une dégradation de l'effort du salarié : $(\underline{w}, \underline{e})$. Cette situation correspond à la pire éventualité du point de vue collectif. Elle est due à l'incitation monétaire de l'employé à dégrader son effort quelle que soit la proposition de salaire qui puisse lui être faite. L'employeur, raisonnant en *induction à rebours* anticipe une telle dégradation de l'effort de son salarié et choisit de ne pas pratiquer le don en proposant le salaire faible.

Cependant compte tenu de la structure des gains du jeu, $(\underline{w}, \underline{e})$ n'est pas le seul candidat à l'équilibre. En effet, il a été supposé que la dégradation de l'effort correspondait à une valeur nulle de ce dernier, procurant un gain nul pour l'employeur quelle que soit sa proposition. Dans cette situation, l'employeur serait indifférent entre proposer le salaire élevé ou le salaire faible même s'il sait que l'employé dégrade son effort en toute éventualité. (\bar{w}, \underline{e}) peut donc constituer un équilibre plausible de ce jeu.

Les éléments de preuve présentés dans l'encadré ci-dessous permettent d'écarter un tel équilibre au profit de $(\underline{w}, \underline{e})$.

Encadré 2-1 : Sélection de l'équilibre de Nash du jeu

Si l'on se place au voisinage de 0 par valeur supérieure pour \underline{e} ; sachant que $\bar{w} > \underline{w}$, et $q = \bar{q}$, constant tout au long du jeu, alors :

$$(q - \bar{w}) < (q - \underline{w}) \text{ implique } \underline{e}(q - \bar{w}) < \underline{e}(q - \underline{w}).$$

Pour toute valeur de \underline{e} tendant vers 0, l'employeur préfère proposer le salaire faible s'il sait que l'employé dégrade son effort. Ce résultat peut être prolongé au cas particulier où $\underline{e} = 0$.

L'employeur anticipant une dégradation de l'effort de l'employé n'est pas indifférent entre proposer le salaire élevé ou le salaire faible; son gain matériel est supérieur dans la seconde éventualité.

L'introduction d'une étape supplémentaire dans ce jeu, visant à permettre à l'employeur de récompenser ou de sanctionner l'employé après avoir observé le niveau d'effort ne perturbe

pas, théoriquement, l'issue non coopérative du jeu. En effet, si ces deux actions sont coûteuses pour l'employeur⁸⁸, leur mise en œuvre n'est pas crédible. Par induction à rebours, le choix du niveau d'effort de l'employé *égoïste* n'est pas perturbé par cette étape supplémentaire du jeu. L'équilibre de Nash, impliquant le choix d'un effort faible et l'absence du don, est donc robuste (voir démonstration en Annexe 2-1).

Etant donné l'équilibre de Nash de ce jeu, il importe d'analyser dans quelle mesure les considérations de réciprocité de l'employeur et de l'employé sont de nature à restaurer l'efficacité des pratiques de don contre don. La prise en compte par l'employeur de telles préoccupations de la part de son employé suffit-elle à le rendre moins méfiant vis à vis de son salarié et à proposer en premier le don sous forme de prime de salaire ? Quelles sont les conditions sur la motivation intrinsèque pour la réciprocité de l'employé ?

Les deux paragraphes suivants répondent à la question de la nature auto-enforçante des pratiques d'échange de cadeau lorsque l'hypothèse de comportements uniquement déterminés par des considérations matérielles est relâchée.

3. COMPORTEMENTS D'ÉQUILIBRE DE L'EMPLOYEUR ET DE L'EMPLOYÉ

L'hypothèse d'*égoïsme* des joueurs est à présent relâchée. Compte tenu des fonctions d'utilité retenues, des conditions sur le paramètre exogène de motivation pour la réciprocité doivent être déterminées pour que, d'une part, l'employeur propose le don, et, d'autre part, l'employé maximise son effort en retour. En d'autres termes, existe-t-il une valeur minimale de ce paramètre telle que la mise en comparaison des conséquences matérielles des actions des joueurs cesse d'être déterminante dans l'adoption des comportements ?

La prise de décision dans ce jeu étant séquentielle, il convient d'abord de montrer qu'une motivation pour la réciprocité suffisante exprimée par l'employé conduit sans ambiguïté à l'adoption d'un comportement coopératif face à la mise en place du don. Si cette condition est remplie, il est possible d'analyser les réactions de l'employeur sachant qu'il n'a plus affaire à un type unique d'employé.

⁸⁸ L'employeur subit un coût de mise en œuvre de la technique de contrôle de l'employé, additionné, le cas échéant du coût direct lié à la récompense allouée.

3.1. Le comportement de l'employé

Les considérations de l'employé pour la réciprocité sont incorporées, comme il a été suggéré dans la première section, par le biais de la fonction d'utilité modifiée proposée par Dufwenberg & Kirchsteiger (1998), particulièrement adaptée au cas des jeux en forme extensive. La partie *psychologique* des fonctions d'utilité est composée de trois éléments tenant compte respectivement de la nature plus ou moins égoïste du joueur (paramètre exogène), d'une *anticipation/observation* sur la bienveillance de son partenaire fondée sur une croyance sur les intentions de celui-ci et enfin d'une *réaction* (vengeance ou récompense).

La proposition suivante est faite :

Proposition 1 : Quand l'employeur propose un salaire faible, l'employé choisit systématiquement de dégrader son effort, qu'il soit motivé ou non par des considérations de réciprocité.

En effet, l'employé reçoit systématiquement un gain inférieur quel que soit son niveau d'effort s'il reçoit \underline{w} (dès lors que $\bar{w} > \underline{w}$ et que $C(e)$ est croissant avec e). Les stratégies $\{(\underline{w}, \bar{e}), (\underline{w}, \underline{e})\}$ sont toujours associées à un gain inférieur comparées aux stratégies $\{(\bar{w}, \bar{e}), (\bar{w}, \underline{e})\}$. Donc même si l'employé est réciproque, une offre \underline{w} est systématiquement interprétée comme une action malveillante de la part de l'employeur. Elle se traduit par une dégradation de l'effort de telle sorte que les deux joueurs aboutissent à l'équilibre de Nash du jeu.

A présent, il convient de s'intéresser au comportement de l'employé lorsqu'il se voit proposer la prime de salaire \bar{w} .

3.1.1. Ecriture de la fonction d'utilité de l'employé recevant \bar{w}

L'employé, confronté à l'offre de l'employeur, juge de la bienveillance de l'action de ce dernier en fonction des intentions qu'il lui attribue et du caractère équitable de la proposition de salaire par rapport à l'ensemble des offres possibles. L'employé choisit alors la réaction appropriée telle qu'elle lui permette de maximiser son utilité.

Bienveillance de l'employé (*réaction*)

La nature bienveillante ou malveillante du comportement du salarié est déterminée par la comparaison entre le gain qu'il permet à l'employeur d'obtenir par son action et le gain qu'il serait juste de procurer à ce dernier (gain équitable), sachant que l'offre correspond à l'hypothèse de salaire haute. Dans ce cas de figure, l'employé peut, dans le choix de son niveau d'effort, procurer au moins 0 et au plus $\bar{e}(q - \bar{w})$ à l'employeur. Le gain équitable auquel l'employeur peut légitimement prétendre est donc :

$$\pi_E^{eq} = \frac{1}{2} \bar{e}(q - \bar{w})$$

La bienveillance de l'employé dans chacune des modalités d'effort s'écrit comme la différence entre le gain de l'employeur et le gain équitable :

S'il maximise son effort :

$$\bar{\kappa}_{LE}^e = \frac{1}{2} \bar{e}(q - \bar{w})$$

S'il dégrade son effort (sachant $\bar{e}=0$) :

$$\bar{\kappa}_{LE}^e = -\frac{1}{2} \bar{e}(q - \bar{w})$$

Compte tenu de ces deux capacités de réaction, il reste à exprimer l'élément qui conditionne cette réaction : l'anticipation de la bienveillance de l'employeur.

Bienveillance de l'employeur (*anticipation/observation*)

Pour juger de la bienveillance ou de la malveillance de l'offre de salaire \bar{w} , l'employé anticipe les intentions qui ont motivé une telle offre. Les intentions attribuées à l'employeur viennent pondérer le caractère objectivement bienveillant ou malveillant de l'offre de salaire. Ainsi, l'adoption de comportements non coopératifs de la part de certains employés réciproques peuvent subsister même si l'offre de salaire est compatible avec la norme considérée comme juste par ceux-ci. Pour ce type d'employé, ce n'est pas seulement le niveau de l'offre qui importe, mais également le sentiment que l'employeur approuve le caractère juste de la norme qu'ils ont définie. En d'autres termes, si l'employé est persuadé de recevoir l'offre de salaire élevée uniquement parce qu'elle constitue, pour l'employeur, le seul moyen de terminer le jeu avec un gain strictement positif, le choix du niveau d'effort peut se trouver

modifié. Cette modélisation donne la possibilité aux joueurs de sanctionner les comportements stratégiques de coopération⁸⁹.

L'employé anticipe donc la croyance *a priori* de l'employeur sur les conséquences de l'offre du don sur l'effort choisi. Cette croyance de croyances ou, plus précisément, cette croyance de deuxième ordre sur la probabilité de choisir l'effort maximum est notée ρ'' ⁹⁰. Ainsi, la croyance de l'employé sur le gain que l'employeur a réellement l'intention de lui procurer dépend de la stratégie effectivement sélectionnée par l'employé entre dégradation et maximisation de son effort, c'est à dire : $\rho''(\bar{w} - C(\bar{e})) + (1 - \rho'')\bar{w}$.

A l'instar de l'expression de sa propre bienveillance, l'anticipation de l'employé sur la bienveillance de l'employeur quand il offre \bar{w} ($\lambda_{LEL}^{\bar{w}}$) peut être définie comme la différence entre le gain qu'il pense recevoir de l'employeur et le gain équitable correspondant à sa prétention légitime compte tenu de la structure du jeu. Dès lors que le gain associé à une proposition \underline{w} est \underline{w} de façon certaine, $\lambda_{LEL}^{\bar{w}}$ peut être exprimée ainsi :

$$\lambda_{LEL}^{\bar{w}} = \frac{1}{2}[\bar{w} - \underline{w} - \rho''C(\bar{e})]$$

L'expression de cette anticipation est intéressante dans le sens où plus l'employeur est persuadé d'une maximisation de l'effort de l'employé, moins il est considéré comme bienveillant par ce dernier. En effet, des employés naturellement *récioproques* considèrent que \bar{w} correspond au *salairre juste (fair wage)* parce qu'ils savent qu'ils maximiseront leur effort en pareille circonstance. Ils ne considèrent donc pas l'employeur comme particulièrement bienveillant lorsqu'il propose une telle offre. Des employés plus ambigus quant à leur motivation pour la réciprocité vont par contre considérer une telle offre comme un réel cadeau et seront plus incités à coopérer malgré leur nature.

⁸⁹ De telles réactions négatives vis à vis de comportements coopératifs stratégiques sont observées dans les expérimentations, notamment dans celle de Thaler (1995). Il montre que les sujets avaient tendance à punir des comportements opportunistes de certains offreurs visant à proposer des partages positifs uniquement pour éviter le refus de la part des répondants et non par soucis de justice distributive.

⁹⁰ L'ajout de la double apostrophe « '' » à la probabilité induit qu'il s'agit de croyances *a priori*. A l'équilibre, l'expression de la probabilité perd ce caractère puisque l'on suppose que les comportements sont compatibles avec les croyances. L'expression de ρ devient alors la probabilité pour l'employé de choisir l'effort maximum.

L'utilité de l'employé, correspondant aux deux décisions d'effort qu'il peut effectuer s'écrit donc :

- s'il maximise son effort :

$$U_L^{\bar{e}} = \bar{w} - C(\bar{e}) + \frac{1}{4} Y_L \bar{e} (q - \bar{w})(\bar{w} - \underline{w} - \rho'' C(\bar{e}))$$

- s'il dégrade son effort :

$$U_L^{\underline{e}} = \bar{w} - \frac{1}{4} Y_L \bar{e} (q - \bar{w})(\bar{w} - \underline{w} - \rho'' C(\bar{e}))$$

A présent, il convient d'analyser les conditions sur la motivation pour la réciprocité de l'employé afin de s'assurer de son comportement coopératif s'il reçoit l'offre de salaire élevée.

3.1.2. Comportement d'équilibre de l'employé selon son type

Etant donnée l'utilité de l'employé selon le choix d'effort qu'il effectue, la condition pour qu'il décide de maximiser son effort est telle que l'utilité associée à une telle stratégie soit supérieure à celle de la paresse, c'est à dire $U_L^{\bar{e}} > U_L^{\underline{e}}$. On obtient donc :

$$\frac{1}{2} Y_L \bar{e} (q - \bar{w})(\bar{w} - \underline{w} - \rho'' C(\bar{e})) > C(\bar{e}) \quad (1)$$

Deux cas sont envisageables à partir de la condition (1) selon le signe du terme $(\bar{w} - \underline{w} - \rho'' C(\bar{e}))$. En effet, selon la différence entre le salaire élevé et le salaire faible, si ρ'' tend vers 1, le terme peut devenir négatif pour un niveau suffisant de coût lié à l'effort. Une telle situation ne peut éventuellement se produire que si la croyance du deuxième ordre est proche de 1 et le coût de l'effort maximum très élevé car sinon, cela voudrait dire que l'offre de salaire élevée ne constituerait pas un réel don par rapport à la situation initiale. Si $(\bar{w} - \underline{w} - \rho'' C(\bar{e})) < 0$ quel que soit ρ'' , le coût associé à l'effort maximum serait tellement important qu'il annihilerait complètement l'avantage associé à l'offre de salaire élevé. Pour que l'offre de salaire élevée soit considérée comme juste, il semble plus plausible d'associer une relation inverse entre la croyance a priori et le niveau du coût de l'effort. En effet, comment penser que l'employé puisse attribuer une croyance a priori élevée, c'est à dire croire que l'employeur s'attend à une maximisation de l'effort, si le coût lié à cette stratégie est très

élevé ? Dans un tel cas de figure l'employeur peut légitimement mettre en doute la volonté de coopération de l'employé puisque celui-ci supporte un coût important à maximiser son effort. De plus, ce dernier est à même d'anticiper un tel sentiment de la part de l'employeur, aussi peut-on s'attendre à ce qu'il associe une faible croyance a priori à l'existence d'un coût de l'effort très important.

En ce qui concerne l'étude du comportement de l'employé, le terme $(\bar{w} - \underline{w} - \rho'' C(\bar{e}))$ est supposé positif. Cette quantité peut être considérée comme le *bénéfice réel subjectif* lié à l'offre de salaire élevé. Cela n'implique pas que le *bénéfice réel objectif* $(\bar{w} - \underline{w} - C(\bar{e}))$ de cette offre soit systématiquement positif. L'importance donnée aux intentions dans le calcul économique des agents implique qu'une offre ne satisfaisant pas la condition de bénéfice réel objectif positif peut malgré tout être considérée comme bienveillante. Pour un employé réciproque, le salaire élevé n'est pas seulement vecteur de bénéfice matériel, il est aussi associé à un statut : la reconnaissance de sa profitabilité par l'employeur. Dans ce cas, ce n'est pas le bénéfice réel qui importe mais la différence entre les deux offres de salaire elle-même⁹¹.

$$(1) \text{ devient alors } Y_L (\bar{w} - \underline{w} - \rho'' C(\bar{e})) > \frac{2C(\bar{e})}{e(q - \bar{w})} \quad (1')$$

Quand l'employeur et l'employé sont à l'équilibre, leurs croyances sont avérées. Si l'employé choisit de maximiser son effort, l'inégalité (1') doit être vérifiée pour $\rho''=1$. La condition pour qu'il coopère implique que sa motivation minimale pour la réciprocité soit telle que :

$$Y_L > \frac{2C(\bar{e})}{e(q - \bar{w})(\bar{w} - \underline{w} - C(\bar{e}))} = Y_L^* > 0.$$

Sa coopération totale est donc conditionnée par une motivation minimale pour la réciprocité. Ce résultat confirme l'idée que la *bienveillance réciproque* agirait comme un mécanisme *d'auto réalisation*⁹² de la coopération dans la relation d'emploi, justifiant la rationalité des pratiques de don contre don observées sur le marché du travail. L'employé suffisamment motivé par la réciprocité répondra favorablement à la pratique du don, faisant passer ses

⁹¹ Les simulations du Chapitre 3 permettent de mettre en évidence la sensibilité du comportement de l'employé vis à vis de la différence entre les deux offres de salaire, toutes choses égales par ailleurs. Il apparaît qu'il existe une différence optimale entre les deux modalités de salaire de sorte que la coopération de l'employé soit maximisée.

⁹² Traduction de *self enforcement*.

considérations matérielles au second plan. Il est prêt à supporter le coût de la mise en œuvre d'un effort supérieur pour récompenser l'offre bienveillante de l'employeur.

L'inégalité (1') ne doit par contre pas être vérifiée pour $\rho^*=0$, cas où l'employé décide de dégrader son effort. Cela se produit si :

$$Y_L < \frac{2C(\bar{e})}{e(q-\bar{w})(\bar{w}-\underline{w})} = Y_L^{**} > 0, \text{ c'est-à-dire s'il est relativement plus attiré par la}$$

maximisation de son gain matériel que par la réciprocité. Dans ce cas, il dégrade systématiquement son effort, même si l'employeur propose la rémunération élevée⁹³. Le coût d'opportunité lié à la maximisation de l'effort est supérieur au supplément d'utilité associé au choix de récompenser une action bienveillante de la part de l'employeur.

Pour des valeurs intermédiaires des seuils entre Y_L^{**} et Y_L^* , un équilibre mixte peut exister, rendant l'employé indifférent entre coopérer ou dégrader son effort quel que soit son type, quand $U_L^{\bar{e}} = U_L^e$. Etant donné la préférence de l'employé pour la réciprocité, sa probabilité de choisir un effort élevé est telle que :

$$\rho = P(e = \bar{e} / w = \bar{w}) = \frac{\bar{w} - \underline{w}}{C(\bar{e})} - \frac{2}{Y_L e(q - \bar{w})} \quad 94.$$

Le premier membre de cette fonction de probabilité représente le gain relatif de l'employé, à se voir proposer le salaire élevé, exprimé en unité d'effort. Dans le second terme, le paramètre de motivation pour la réciprocité intervient de façon positive sur la probabilité de coopération de l'employé. En effet, plus le type de l'employé se rapproche du seuil correspondant aux joueurs *réiproques*, plus la proposition \bar{w} est considérée comme bienveillante. A mesure que le type de l'employé se rapproche du seuil Y_L^* , la satisfaction ressentie par l'employé à récompenser l'action bienveillante de l'employeur dépasse la désutilité de l'effort supplémentaire mis en œuvre.

La palette des comportements possibles de l'employé en fonction de son type est présentée dans le tableau suivant :

⁹³ On vérifie que $Y_L^{**} < Y_L^*$ dès lors que $C(\bar{e}) > 0$.

⁹⁴ On peut vérifier que pour les employés dont la préférence intrinsèque pour la réciprocité est Y_L^{**} , cette probabilité est nulle. En revanche elle est égale à 1 pour les employés de type Y_L^* . On vérifie également que cette probabilité est monotone et croissante sur l'intervalle décrit par ces deux seuils.

Tableau 2-2 : Comportement d'équilibre de l'employé selon son type et la proposition de salaire.

	Valeur de Y_L (motivation pour la réciprocité)		
Proposition, de salaire	0	Y_L^{**}	Y_L^*
\underline{w}	Dégradation de l'effort : \underline{e}		
\bar{w}	\underline{e}	$P(e=\bar{e}) = \rho$	\bar{e}

Ces résultats corroborent l'idée que la motivation pour la *bienveillance réciproque* agit comme condition nécessaire à l'établissement de pratiques de don contre don réussies entre employeur et employé. Si ce dernier accorde suffisamment d'importance à de telles considérations, sa satisfaction totale à maximiser son effort dans de telles circonstances surpasse la satisfaction monétaire qu'il pourrait retirer d'une dégradation de son effort. Le coût d'opportunité associé à la maximisation de son effort est plus faible que le coût psychologique qu'il ressent à être injuste vis à vis d'un partenaire bienveillant (sentiment de culpabilité). Le résultat de cette comparaison est directement lié à sa nature plus ou moins égoïste qui tend à modifier la pondération qu'il accorde à ses deux motivations opposées. On retrouve ici les écarts observés, par exemple, dans le jeu du dictateur où une offre généreuse de la part du joueur **X** est interprétée comme le résultat d'une telle comparaison entre appât du gain et sentiment de culpabilité à extorquer une part importante d'un surplus exogène. Il est ainsi possible d'énoncer la proposition suivante concernant le comportement d'équilibre du salarié :

Proposition 2 : Un employé suffisamment motivé par la réciprocité retourne le don de l'employeur (sous forme de prime de salaire) par une maximisation de son effort. La réciprocité agit comme condition nécessaire et suffisante, du point de vue de l'employé, pour induire des comportements coopératifs.

Cependant, l'employé prend sa décision après avoir observé la proposition de salaire. L'incertitude porte seulement, de son point de vue, sur l'intention qu'il attribue à l'employeur lorsque celui-ci propose un salaire élevé. Quant à l'employeur, il prend sa décision avant même d'observer le comportement de son employé. Une telle asymétrie de position entre employé et employeur pose la question de la capacité de confiance de ce dernier, à l'instar de l'offreur dans le jeu d'investissement, et peut remettre en cause sa volonté de pratiquer le don

en premier. La structure même du jeu peut donc conduire un employeur motivé par la réciprocité à se considérer comme pris en otage dans la relation et à remettre en cause le caractère juste d'une proposition de salaire élevée. Le paragraphe suivant présente l'analyse du comportement de l'employeur afin de déterminer si de tels arguments sont susceptibles de menacer l'apparition d'échanges de cadeaux entre les deux joueurs.

3.2. Le comportement de l'employeur

L'employeur peut interagir avec trois types possibles d'employés : les égoïstes ($Y_L < Y_L^{**}$), les *réiproques* ($Y_L > Y_L^*$), et enfin les employés dont la motivation est plus ambiguë ($Y_L^{**} < Y_L < Y_L^*$). Chacun des cas est envisagé dans l'écriture de la fonction d'utilité.

3.2.1. Ecriture de la fonction d'utilité de l'employeur

La proposition suivante concerne le comportement d'équilibre de l'employeur lorsqu'il est apparié avec un employé égoïste. Ensuite chacun des autres cas est envisagé.

Proposition 3 : Pour $Y_L < Y_L^{**} = \frac{2C(\bar{e})}{e(q - \bar{w})(\bar{w} - w)}$, l'employeur sait que l'employé est motivé

principalement par la maximisation de son gain monétaire et donc qu'il va dégrader son effort, quelle que soit l'offre. Dès lors qu'il anticipe un gain nul, il préfère proposer le salaire faible, quel que soit son type. L'équilibre est donc $(\underline{w}, \underline{e})$, l'équilibre de Nash du jeu.

3.2.1.1. Si l'employeur est apparié avec un employé réiproque : $Y_L > Y_L^*$

Bienveillance de l'employeur

Comme pour l'employé, la bienveillance ou malveillance de l'action de l'employeur est déterminée par la comparaison entre le gain qu'il procure au salarié par sa proposition et le gain équitable que ce dernier est en droit d'attendre à l'issue de la relation. Il peut lui procurer

au minimum \underline{w} , et au maximum $\bar{w} - C(\bar{e})$. La proposition correspondant à une offre juste est donc :

$$\pi_L^{eq} = \frac{1}{2}(\bar{w} + \underline{w} - C(\bar{e}))$$

La bienveillance de l'employeur, pour chacune des proposition est définie ainsi :

- S'il propose le salaire élevé :

$$K_{EL}^{\bar{w}} = \frac{1}{2}(\bar{w} - \underline{w} - C(\bar{e}))$$

- S'il propose le salaire faible :

$$K_{EL}^{\underline{w}} = -\frac{1}{2}(\bar{w} - \underline{w} - C(\bar{e}))$$

Comme pour l'employé, il faut déterminer son anticipation sur la bienveillance de ce dernier pour écrire sa fonction d'utilité.

Bienveillance anticipée de l'employé

Pour juger de la bienveillance de l'employé, l'employeur anticipe ses intentions de comportement face à l'offre qui lui est faite. Pour cela, il anticipe la croyance a priori de l'employé sur la proposition qui va effectivement lui parvenir. Cette croyance de deuxième ordre sur la probabilité de choisir d'offrir la prime de salaire est notée ε'' . Dans ce cas, l'employeur pense que l'employé lui procurera $\varepsilon'' \bar{e}(q - \bar{w}) + (1 - \varepsilon'') 0$ en choisissant sa stratégie d'équilibre $\{(\bar{w}, \bar{e}), (\underline{w}, \underline{e})\}$.

Si l'employé choisit toujours de maximiser son effort quelle que soit l'offre de salaire, l'employeur peut recevoir un gain matériel de $\varepsilon'' \bar{e}(q - \bar{w}) + (1 - \varepsilon'') \bar{e}(q - \underline{w})$. Au contraire, s'il pense que l'employé dégradera toujours son effort, l'employeur reçoit un gain nul. De ce point de vue, le *gain équitable* que l'employeur peut légitimement attendre de la relation est donc défini par :

$$\pi_E^{eq} = \frac{1}{2}[\varepsilon'' \bar{e}(q - \bar{w}) + (1 - \varepsilon'') \bar{e}(q - \underline{w}) + 0]$$

La croyance de l'employeur sur la bienveillance de l'employé quand il dégrade son effort face à une proposition de salaire faible et maximise son effort quand il reçoit \bar{w} , notée $\lambda_{ELE}^{\{(w,e);(\bar{w},\bar{e})\}}$ s'écrit de la même façon que celle de l'employé. Elle correspond à la différence entre le gain

qui échoit à l'employeur lorsque son employé est réciproque et le gain équitable exprimé plus haut, sachant que l'on se trouve à présent au premier nœud du jeu⁹⁵, d'où :

$$\lambda_{ELE}^{\{(w,e);(\bar{w},\bar{e})\}} = \frac{1}{2}\varepsilon''\bar{e}(q-\bar{w}) - \frac{1}{2}(1-\varepsilon'')\bar{e}(q-\underline{w})$$

L'utilité de l'employeur pour chaque offre de salaire qu'il peut proposer s'écrit :

- s'il propose le don (salaire élevé) :

$$U_E^{\bar{w}} = \pi_E(\bar{w}, \bar{e}) + Y_E K_{EL}^{\bar{w}} \lambda_{ELE}^{\{(\bar{w}, \bar{e}); (w, e)\}}, \text{ c'est-à-dire :}$$

$$U_E^{\bar{w}} = \varepsilon''\bar{e}(q-\bar{w}) + \frac{1}{4}Y_E \left[\varepsilon''\bar{e}(q-\bar{w})(\bar{w}-\underline{w}-C(\bar{e})) - (1-\varepsilon'')\bar{e}(q-\underline{w})(\bar{w}-\underline{w}-C(\bar{e})) \right]$$

- s'il propose le salaire faible :

$$U_E^{\underline{w}} = \pi_E(\underline{w}, \underline{e}) + Y_E K_{EL}^{\underline{w}} \lambda_{ELE}^{\{(\bar{w}, \bar{e}); (w, e)\}}, \text{ c'est-à-dire :}$$

$$U_E^{\underline{w}} = -\frac{1}{4}Y_E \left[\varepsilon''\bar{e}(q-\bar{w})(\bar{w}-\underline{w}-C(\bar{e})) - (1-\varepsilon'')\bar{e}(q-\underline{w})(\bar{w}-\underline{w}-C(\bar{e})) \right]$$

3.2.1.2. Si l'employeur est apparié avec un employé ambigu quant à ses motivations. $Y_L^{**} < Y_L < Y_L^*$

De la même façon que lorsque l'employeur est apparié avec un employé réciproque, son utilité associée à chaque proposition de salaire peut être écrite ainsi (Cf. démonstration en Annexe 2-2) :

- s'il propose le don (salaire élevé) :

$$U_E^{\bar{w}} = \pi_E(\bar{w}, p(e=\bar{e})=\rho) + Y_E K_{EL}^{\bar{w}} \lambda_{ELE}^{\{(w,e);(\bar{w}, p(e=\bar{e})=\rho)\}}, \text{ c'est-à-dire :}$$

$$U_E^{\bar{w}} = \rho\bar{e}(q-\bar{w}) + \frac{1}{2}Y_E(\bar{w}-\underline{w}-\rho C(\bar{e})) \left[\varepsilon''\bar{e}(q-\bar{w})(\rho-\frac{1}{2}) - \frac{1}{2}(1-\varepsilon'')\bar{e}(q-\underline{w}) \right]$$

- s'il propose le salaire faible :

$$U_E^{\underline{w}} = \pi_E(\underline{w}, \underline{e}) + Y_E K_{EL}^{\underline{w}} \lambda_{ELE}^{\{(w,e);(\bar{w}, p(e=\bar{e})=\rho)\}}, \text{ c'est-à-dire :}$$

$$U_E^{\underline{w}} = -\frac{1}{2}Y_E(\bar{w}-\underline{w}-\rho C(\bar{e})) \left[\varepsilon''\bar{e}(q-\bar{w})(\rho-\frac{1}{2}) - \frac{1}{2}(1-\varepsilon'')\bar{e}(q-\underline{w}) \right]$$

⁹⁵ Le concept de *Sequential Reciprocity Equilibrium* développé par Dufwenberg & Kirchsteiger (1998) inclut la possibilité pour les joueurs de réviser leurs croyances au cours du jeu. Ces révisions induisent qu'à chaque nœud du jeu correspond un gain équitable différent et, par conséquent, des croyances différentes sur la bienveillance de l'autre.

(rappel : l'employé coopère avec la probabilité ρ dans ce cas, ce qui modifie substantiellement l'anticipation de l'employeur sur sa bienveillance).

Par la suite, dans l'écriture des conditions sur le paramètre de motivation pour la réciprocité de l'employeur, seulement le cas où il est apparié avec un employé réciproque est considéré. Cette présentation est motivée par un souci de clarté de la démonstration. Le lecteur est invité à se reporter à l'Annexe 2-3 pour la démonstration du comportement d'équilibre de l'employeur apparié avec un employé modérément motivé par la réciprocité.

3.2.2. Comportement d'équilibre de l'employeur selon le type d'employé

La condition pour que l'employeur propose la prime de salaire est déterminée par la comparaison de l'utilité associée aux deux offres possibles, avec $U_E^{\bar{w}} > U_E^{\underline{w}}$, ce qui implique :

$$\bar{e}(q - \bar{w}) + \frac{1}{2} Y_E (\bar{w} - \underline{w} - C(\bar{e})) (\varepsilon'' \bar{e}(q - \bar{w}) - (1 - \varepsilon'') \bar{e}(q - \underline{w})) > 0 \quad (2)$$

La recherche des conditions sur la motivation propre de l'employeur à privilégier la justice réciproque impose une distinction des comportements possibles suivant le signe de l'élément $(\bar{w} - \underline{w} - C(\bar{e}))$. En effet, si cette quantité est négative, cela signifie que le bénéfice réel (objectif) que retire le salarié de la prime de salaire devient négatif. Dans quelle mesure, face à une telle situation, l'employeur peut-il raisonnablement s'attendre à une maximisation de l'effort de son salarié ?

De tels comportements coopératifs ne peuvent provenir que d'employés fortement motivés par des considérations de bienveillance (*fairness*) et non d'employés stratégiques. Cependant, l'analyse du comportement de l'employé montre que son choix est conditionné par le jugement favorable qu'il peut déduire de l'action de l'employeur. Pour qu'il coopère, dans cette situation de désavantage objectif, à accepter la forte rémunération $(\bar{w} - \underline{w} - C(\bar{e}) < 0)$, un autre élément doit représenter pour lui une source de contentement dans cette proposition de salaire. Cet élément doit être de plus de connaissance commune. Ceci sous-tend l'idée que pour certains employés ce n'est pas tant la valeur associée à \bar{w} qui importe mais le statut qui lui est attaché. Ainsi, des salariés à fort potentiel du point de vue de l'employeur (q élevé) peuvent trouver un intérêt à se voir reconnaître une telle qualité par l'octroi d'un niveau de salaire élevé. Celui-ci constitue un signal de leur profitabilité, même si pour cela le coût de la

maximisation de l'effort dépasse l'ampleur de la promotion salariale. La proposition \bar{w} devient bienveillante du point de vue de tels employés « réciproques » et motive leur coopération. Le marché du travail fournit de nombreuses illustrations de ce phénomène dont on peut trouver un exemple caricatural sur le marché de la grande distribution. Les employés des supermarchés sont rémunérés sur la base d'un volume horaire hebdomadaire prévu dans le contrat. Lorsqu'ils se voient proposer le statut de chef de rayon, assimilé à un emploi de cadre, le volume horaire n'est plus incorporé dans le contrat si bien que le nombre d'heures de présence sur le lieu de travail augmente considérablement. Le coût de l'effort associé à ces nouvelles conditions d'exécution du contrat dépasse alors souvent l'attrait financier de la promotion. La proposition est pourtant généralement considérée comme bienveillante car elle est interprétée comme une reconnaissance de la valeur de l'employé. On peut donc dire que le cas $\bar{w} - \underline{w} - C(\bar{e}) < 0$ peut s'appliquer à des offres de salaires induisant un changement important dans la relation, comme l'accès à un emploi de cadre. Dans le modèle, cet effet permet d'introduire une nuance entre valeur du bénéfice réel ($\bar{w} - \underline{w} - C(\bar{e})$) et valeur du bénéfice nominal ($\bar{w} - \underline{w}$) lié à l'offre de salaire élevé. En effet, l'employé évalue la bienveillance de l'offre à travers la fonction : $\lambda_{LEL}^{\bar{w}} = \frac{1}{2}[\bar{w} - \underline{w} - \rho''C(\bar{e})]$. L'employé fortement motivé par la réciprocité accorde une importance particulière à la valeur de cette fonction, largement déterminée par le paramètre de croyance sur les intentions de l'employeur. Si le bénéfice nominal est positif, l'offre de salaire élevé peut encore être considérée comme bienveillante dès lors que ρ'' est très faible, alors qu'objectivement le bénéfice réel de l'offre est négatif. Ces employés motivés par un besoin de reconnaissance peuvent donc considérer l'offre comme bienveillante dès lors qu'elle leur permet simplement d'accéder à une catégorie de rémunération supérieure. La coopération peut donc émerger d'une telle offre mêlant bénéfice réel négatif et bénéfice nominal positif. Par conséquent, le cas où $\bar{w} - \underline{w} - C(\bar{e})$ est négatif reste pertinent pour l'analyse du comportement d'offre de l'employeur. Il est même possible de poursuivre le raisonnement en remarquant que cet effet bénéfice nominal vs bénéfice réel conduit à mettre en évidence l'existence d'une différence optimale entre les deux offres de salaire dans la détermination des choix de comportement des deux joueurs. Cet effet est observé dans les simulations proposées dans le chapitre suivant (Cf. Chapitre 3).

L'analyse porte à présent sur les deux cas possibles concernant le bénéfice associé à la proposition de salaire élevée :

si $\bar{w} - \underline{w} - C(\bar{e}) > 0$, c'est à dire s'il existe un avantage réel associé à \bar{w} pour l'employé maximisant son effort.

$$(2) \text{ devient } Y_E > -\frac{2(q - \bar{w})}{\varepsilon''(q - \bar{w})(\bar{w} - \underline{w} - C(\bar{e})) - (1 - \varepsilon'')(q - \underline{w})(\bar{w} - \underline{w} - C(\bar{e}))} \quad (2')$$

Quand l'employeur et l'employé sont à l'équilibre, leurs croyances sont avérées. Par conséquent si l'employeur décide d'offrir \bar{w} , la condition (2') doit être vérifiée pour $\varepsilon'' = 1$. Sa motivation pour la réciprocité doit alors être telle que :

$$Y_E > -\frac{2}{(\bar{w} - \underline{w} - C(\bar{e}))} = Y_E^* < 0$$

Sachant que ce paramètre est borné inférieurement en 0, cette condition apparaît comme peu contraignante. Un employeur égoïste apparié à un employé réciproque peut tout à fait proposer la rémunération élevée. Ce résultat paraît relativement logique car un tel employeur peut stratégiquement proposer l'offre \bar{w} puisqu'il sait qu'elle sera assortie d'un gain supérieur pour lui compte tenu de la nature réciproque de l'employé. Un employeur égoïste a intérêt d'exploiter à son profit la motivation sociale de l'employé. L'élément pertinent pour ce type d'employeur devient la croyance de l'employé sur ses intentions⁹⁶.

Si l'employeur décide d'offrir \underline{w} , la condition (2') ne doit pas être remplie pour $\varepsilon'' = 0$, c'est-à-dire pour des valeurs de la motivation intrinsèque pour la réciprocité telles que :

$$Y_E > \frac{2(q - \bar{w})}{(q - \underline{w})(\bar{w} - \underline{w} - C(\bar{e}))} = \tilde{Y}_E^{**} > 0$$

Ce résultat théorique peut paraître contre intuitif dans la mesure où cette condition implique qu'un employeur motivé par la réciprocité aurait tendance à être plus méfiant et à ne pas proposer le don. L'explication possible se trouve dans la structure informationnelle du jeu. L'employeur, contrairement à l'employé, doit prendre sa décision sans pouvoir observer ex ante une quelconque action de l'employé. Il se trouve en position de faiblesse vis à vis de son partenaire et peut éprouver un sentiment de « prise en otage », ce qui affecte son anticipation sur la bienveillance de l'autre. Comme il l'a déjà été souligné précédemment, les intentions attribuées au partenaire déterminent grandement une telle anticipation, l'employeur est

⁹⁶ Ce résultat théorique amène à nuancer l'interprétation des résultats expérimentaux de Fehr & al (1996). Dans leur expérimentation des comportements de réciprocité par le biais d'un jeu d'enchère, ils associent motivation pour la réciprocité et comportement coopératifs de leurs vendeurs. Ce résultat théorique vient souligner l'intérêt de distinguer entre action réciproque et réelle motivation pour la réciprocité.

incapable, à ce nœud précis du jeu, d'évaluer si l'employé se comporte de façon stratégique ou selon des considérations de réciprocité. Il apparaît donc que la « confiance mutuelle » agit comme une contrainte dès lors que les intentions des joueurs ne sont pas clairement identifiables.

De tels comportements correspondent à des équilibres de *prophéties auto réalisatrices* de sorte que l'employeur envisage la pire issue du jeu et ajuste sa décision en conséquence.

si $\bar{w} - \underline{w} - C(\bar{e}) \leq 0$,

$$(2) \text{ devient } Y_E < -\frac{2(q - \bar{w})}{\varepsilon''(q - \bar{w})(\bar{w} - \underline{w} - C(\bar{e})) - (1 - \varepsilon'')(q - \underline{w})(\bar{w} - \underline{w} - C(\bar{e}))} \quad (2'')$$

De la même manière que dans les cas précédents, (2'') doit être vérifiée pour $\varepsilon'' = 1$ si l'employeur décide de proposer \bar{w} , et ne doit pas être remplie pour $\varepsilon'' = 0$ dans l'éventualité contraire. Les conditions sur le paramètre de réciprocité deviennent alors :

$$Y_E < -\frac{2}{(\bar{w} - \underline{w} - C(\bar{e}))} = \tilde{Y}_E^* > 0, \text{ pour } \varepsilon'' = 1, \text{ et}$$

$$Y_E < \frac{2(q - \bar{w})}{(q - \underline{w})(\bar{w} - \underline{w} - C(\bar{e}))} < 0, \text{ pour } \varepsilon'' = 0, \text{ ce qui ne peut être le cas puisque } Y_E \geq 0.$$

Pour les valeurs du paramètre de réciprocité au-delà de \tilde{Y}_E^* , il existe un équilibre mixte rendant l'employeur indifférent entre proposer \bar{w} et \underline{w} . On a $U_E^{\bar{w}} = U_E^{\underline{w}}$, ce qui implique :

$$\varepsilon[w = \bar{w} / \{(\bar{w}, \bar{e}), (\underline{w}, \underline{e})\}] = \frac{q - \underline{w}}{(q - \bar{w}) + (q - \underline{w})} - \frac{2(q - \bar{w})}{Y_E(\bar{w} - \underline{w} - C(\bar{e}))(2q - \bar{w} - \underline{w})}$$

$\varepsilon(\cdot)$ correspond donc à la probabilité que l'employeur propose la rémunération élevée sachant que l'employé est réciproque et adopte sa stratégie d'équilibre correspondant à chaque offre de salaire. Le premier terme représente le gain relatif de l'employeur à proposer la rémunération faible si l'employé coopérait quelle que soit la proposition. Ce terme est pondéré par un deuxième terme tenant compte du bénéfice réel de l'employé associé à \bar{w} (qui est négatif dans le cas présent) et de la motivation de l'employeur pour la réciprocité⁹⁷.

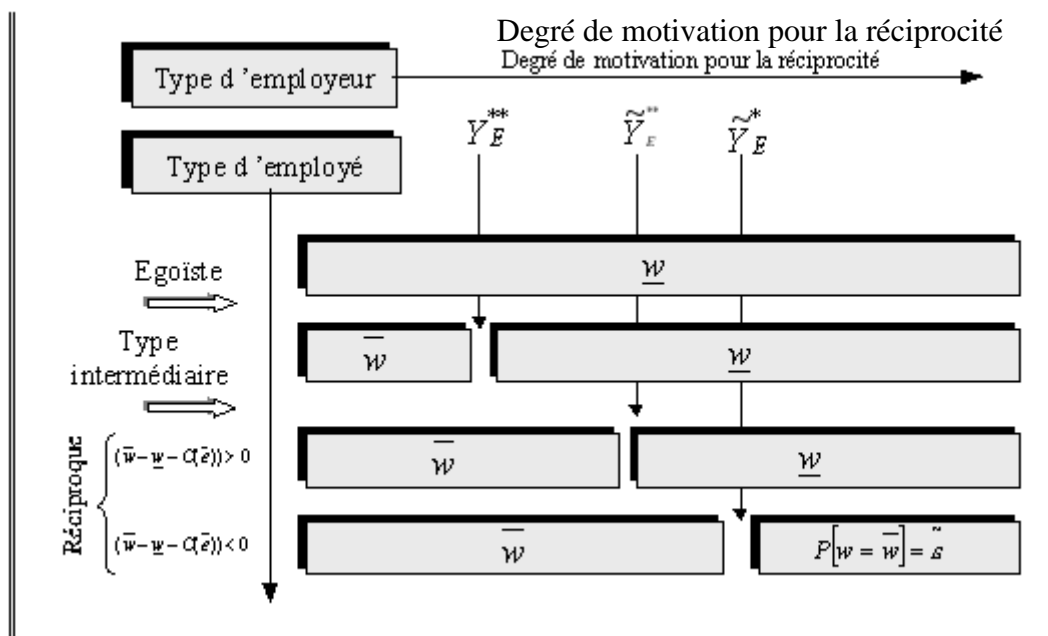
Le signe de $\bar{w} - \underline{w} - C(\bar{e})$ exerce une influence sur le comportement de l'employeur. Si la différence est positive, son comportement n'implique aucune incertitude ; son type, plus ou

⁹⁷ Ce paramètre joue négativement sur la probabilité totale puisque le terme qui le suit est négatif. La probabilité diminue à mesure que l'employeur est motivé par la réciprocité.

moins réciproque, détermine sa décision. Dans le deuxième cas, compte tenu qu'il sait que toute coopération de la part de l'employé ne peut provenir que de joueurs réellement réciproques⁹⁸, il a plus de scrupules à ne pas proposer le don s'il est lui-même réciproque. C'est pourquoi, au-delà de \tilde{Y}_E^* , il ne choisit plus systématiquement de proposer \underline{w} mais laisse la possibilité à une pratique risquée du don avec une probabilité positive.

La tableau suivant permet de résumer le comportement d'équilibre de l'employeur selon son type et selon la nature de l'employé avec lequel il est apparié.

Tableau 2-3 : Comportement d'équilibre de l'employeur selon son type et celui de l'employé.



Dans le cas où l'employé est modérément motivé par la réciprocité, la zone d'offre du salaire élevé est logiquement plus réduite et correspond par conséquent à des employeurs faiblement réciproques, en deçà de $Y_E^{**} = \frac{2\rho(q - \bar{w})}{(q - \underline{w})(\bar{w} - \underline{w} - \rho C(\bar{e}))}$ ⁹⁹.

⁹⁸ Les coopérateurs stratégiques n'ont aucun intérêt à maximiser leur effort si $\bar{w} - \underline{w} - C(\bar{e}) < 0$.

⁹⁹ Le seuil Y_E^{**} diffère de \tilde{Y}_E^{**} , par l'élément ρ qui vient pondérer le numérateur et le dénominateur. Cet élément représente la probabilité pour que l'employé modérément motivé par la réciprocité coopère. (Voir la démonstration en Annexe 3 pour l'étude du comportement d'équilibre de l'employeur lorsqu'il est apparié avec un tel employé).

L'analyse du comportement d'équilibre de l'employeur induit les deux propositions suivantes :

Proposition 4 : Un employeur égoïste apparié à un employé réciproque pratique le don pour des raisons stratégiques de maximisation de son gain final.

L'employeur égoïste anticipe que l'employé réciproque peut être enclin à maximiser son effort s'il offre le salaire élevé. Etant peu motivé par la réciprocité, la partie psychologique de ses gains influence peu la valeur de son utilité. Ce comportement d'offre de salaire élevé stratégique prédomine tant que le type de l'employeur est de telle sorte qu'il valorise ses gains psychologiques en deçà d'une valeur seuil Y_E^{**} . Ce comportement peut être qualifié de coopération stratégique car, bien qu'étant de nature égoïste, l'employeur tient compte de la manière dont est interprétée son offre de salaire élevé par l'employé. En effet, ses gains matériels sont pondérés par le paramètre de croyance de second ordre de l'employé (ρ). Il prend sa décision de pratique du don s'il pense que l'employé interprétera cette offre comme bienveillante. Il s'agit bien d'une offre stratégique puisque l'employeur se fait passer pour un joueur bienveillant alors que la seule motivation qui le pousse à pratiquer le don est d'ordre purement matériel.

Proposition 5 : Le comportement d'équilibre de l'employeur n'est pas unique en fonction de son type et celui de l'employé. Il dépend également de la nature plus ou moins profitable de l'offre de salaire élevée.

Quel que soit son type, l'employeur attache de l'importance aux conséquences matérielles de son offre sur l'employé. S'il est coopérateur stratégique, il s'attache à procurer un gain suffisamment important pour que l'offre soit considérée comme volontairement bienveillante. Dans cette perspective, il restreint son offre de salaire élevé aux seuls cas où le coût de l'effort impliqué par la coopération de l'employé produit un bénéfice réel positif à ce dernier ($\bar{w} - \underline{w} - C(\bar{e}) > 0$). Ici, la variable de décision est représentée par l'anticipation de l'employé sur l'intention de l'employeur. En revanche, si l'employeur est réciproque, la nature réellement bienveillante de son action (\bar{K}_{EL}^w) détermine son choix¹⁰⁰.

Etant donnés les comportements d'équilibre des deux protagonistes à la négociation, il est possible de dresser une typologie des équilibres du jeu et de répondre à la question des conditions d'apparition et d'efficacité des pratiques de don dans la détermination des salaires.

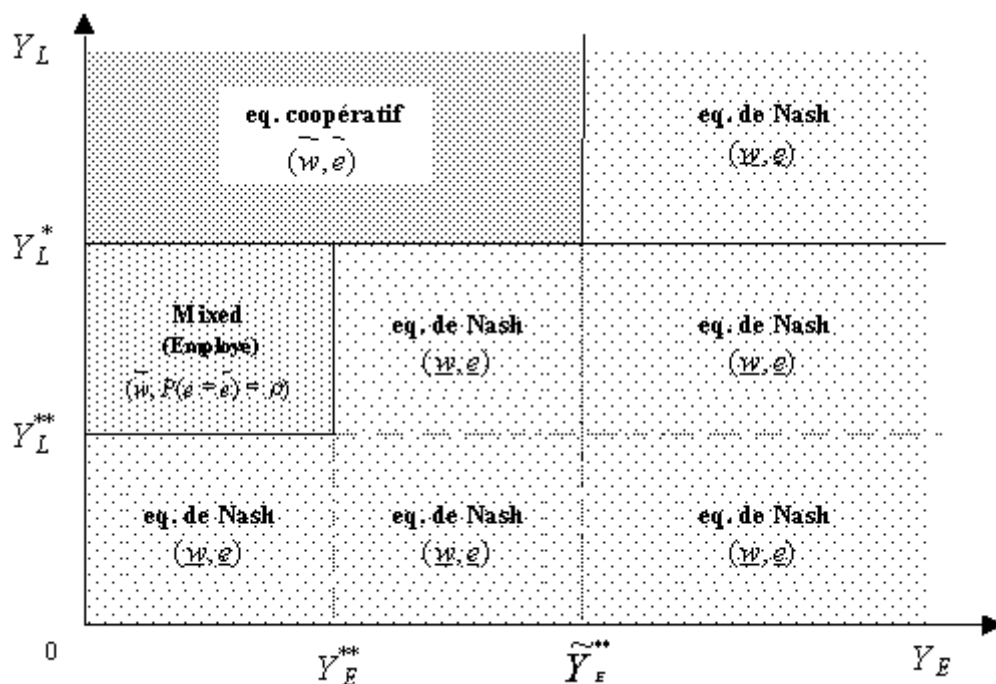
4. NATURE DES ÉQUILIBRES THÉORIQUES ET CONCLUSIONS SUR L'EFFICACITÉ DES PRATIQUES DE DON DANS LA DÉTERMINATION DES SALAIRES

Les deux graphiques suivants expriment les équilibres possibles de la négociation entre employeur et employé issus de la confrontation des comportements d'équilibre de ces deux joueurs exprimés précédemment.

La distinction est opérée sur la base du signe de $(\bar{w} - \underline{w} - C(\bar{e}))$ qui rend l'offre de salaire élevée plus ou moins attractive du point de vue de l'employé.

Si la différence est positive: $(\bar{w} - \underline{w} - C(\bar{e})) > 0$

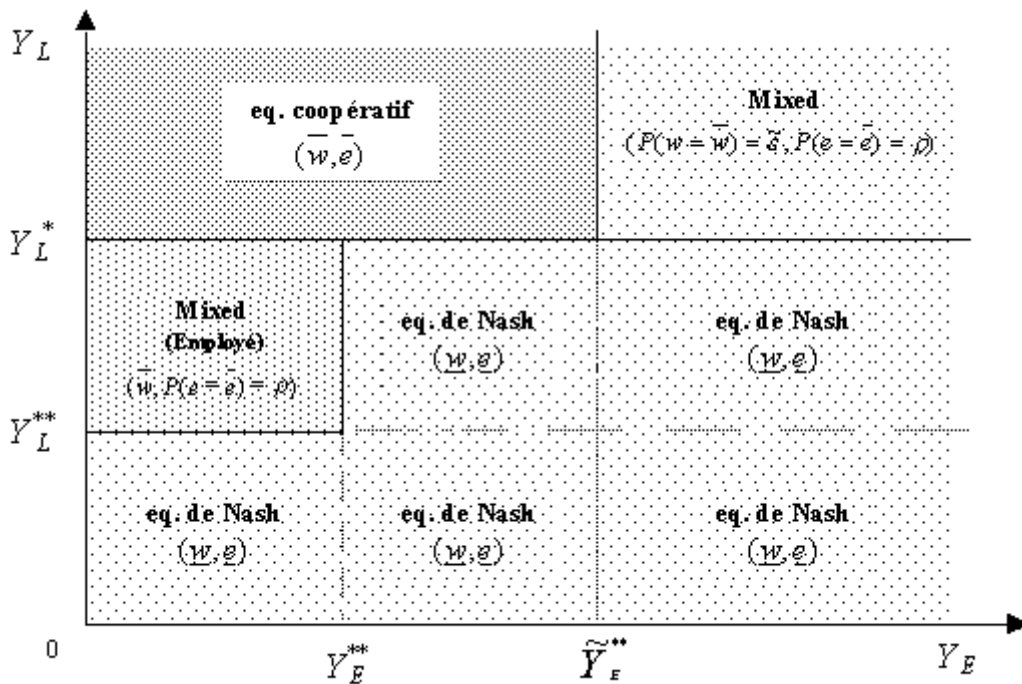
Graph 2-2 : Equilibres du jeu d'échange de cadeaux en fonction des types de joueurs, $(\bar{w} - \underline{w} - C(\bar{e})) > 0$.



¹⁰⁰ Les simulations proposées au chapitre 3 permettent de préciser les valeurs critiques de ces paramètres

- Si la différence est négative: $(\bar{w} - \underline{w} - C(\bar{e})) < 0$

Graphe 2-3 : Equilibres du jeu d'échange de cadeaux en fonction des types de joueurs, $(\bar{w} - \underline{w} - C(\bar{e})) < 0$.



Pour les deux graphiques, en abscisse figure le type de l'employeur; du plus égoïste (proche de 0) au plus réciproque. Les trois types possibles d'employés se trouvent en ordonnée.

Il apparaît clairement que la motivation pour la réciprocité agit comme condition nécessaire et suffisante du point de vue du salarié pour susciter une maximisation de l'effort suite à un éventuel don de la part de l'employeur. L'employé valorisant la réciprocité au-delà d'une valeur Y_L^* retourne le don qui lui est fait par la proposition \bar{w} . D'autre part, ce résultat n'est pas perturbé lorsque le bénéfice réel associé à \bar{w} devient négatif, ce qui confirme l'idée préalable que seuls les employés identifiés comme réellement « réciproques » coopèrent dans une telle éventualité. L'observation de Fehr et al. (1996) d'une corrélation positive entre salaire offert et effort coûteux peut alors être théoriquement attribuée à des considérations de bienveillance réciproque exprimées par l'employé.

Le modèle permet donc de montrer que si le don est pratiqué en présence d'employés sensibles à la bienveillance de l'employeur, l'efficacité d'une telle pratique est assurée par

conditionnant le choix d'offre du don de la part de l'employeur.

l'adoption systématique de la stratégie de maximisation de l'effort. Cependant, les employés expriment une aversion aux comportements d'offre stratégique du don d'autant plus forte que le bénéfice réel associé à celle-ci est faible, c'est-à-dire lorsque le coût lié à l'effort est particulièrement élevé. Des échecs du don réciproque apparaissent si l'intention de l'employeur n'est pas clairement identifiée comme étant bienveillante. De ce point de vue, il existe une sorte d'arbitrage entre *bénéfice réel objectif* de l'offre et intention attribuée à l'employeur dans la détermination de sa bienveillance. Cet élément conduit à suspecter l'existence d'une valeur optimale pour ce bénéfice, et, par voie de conséquence, d'une valeur optimale de la différence entre les deux modalités de salaire, telle que la coopération de l'employé soit maximisée. Ainsi, en ne tenant compte que des considérations de réciprocité dans l'explication des comportements d'échange de cadeaux, les résultats théoriques indiquent que la relation positive entre salaire et effort coûteux atteint un maximum déterminé par l'ampleur de la promotion relative au coût de l'effort. Au-delà de cette valeur optimale, deux éléments concourent à la dégradation de la probabilité de retour du don. D'une part, plus la différence de salaire est élevée, plus la partie matérielle de la fonction d'utilité de l'employé est importante. L'exigence en terme de motivation pour la réciprocité de l'employé est donc plus contraignante (le seuil minimum croît avec la valeur des gains matériels). D'autre part, pour les employés satisfaisant cette exigence, l'attrait de l'offre de salaire élevé a des répercussions négatives sur l'anticipation des intentions de l'employeur. Le chapitre suivant permettra de mettre en évidence l'influence de cette différence optimale et de préciser sa valeur dans les différents contextes envisagés.

Au-delà du comportement de l'employé, les résultats théoriques montrent que la difficulté provient essentiellement de l'apparition réelle du don de la part de l'employeur. En effet, les deux graphiques montrent qu'il existe une plage de stabilité certaine pour une stratégie non coopératrice de la part de l'employeur « réciproque » (Cf. équilibres situés en haut à droite sur les deux graphes). Un tel comportement est alors immédiatement sanctionné par l'employé, quelle que soit sa nature. Il semble que cette absence de pratique du don par ce type d'employeur soit la conséquence d'une anticipation erronée sur la stratégie d'équilibre de l'employé. Il s'ensuit donc l'apparition d'équilibres pouvant être qualifiés « d'équilibres de prophéties auto réalisatrices » qui mettent en danger l'efficacité de la relation d'emploi. Cette inefficience peut être attribuée, comme il l'a été vu précédemment, à la structure asymétrique du jeu concernant l'expression des gains psychologiques des joueurs. L'employé peut déterminer la bienveillance de l'employeur à partir de l'observation de l'offre qui lui est faite. L'employeur, quant à lui, doit évaluer cette bienveillance à travers une anticipation sur le

comportement d'équilibre de l'employé, sans observation tangible d'un comportement passé puisque le jeu est exprimé en un coup. Ainsi, l'équilibre de prophéties auto réalisatrices provient d'une anticipation pessimiste de l'employeur sur la croyance de l'employé. Il pense que le salarié s'attend à une offre de salaire faible, éprouvant une certaine amertume *a priori* vis à vis de lui. L'employeur anticipe alors une stratégie de dégradation systématique de l'effort quelle que soit l'offre, ce qui motive une offre faible. S'il propose le don, il s'attend à ce que son offre soit interprétée comme stratégique.

L'employeur peut à juste titre craindre une telle dégradation de l'effort suite à une proposition \bar{w} . Compte tenu de l'écriture du comportement d'équilibre de l'employeur égoïste (offre systématique de \bar{w} ; Cf. Tableau 2-3, page 149), l'offre \bar{w} peut amener l'employé à se méprendre sur la nature de l'employeur. Si tel est le cas, l'anticipation qu'il forme sur la bienveillance de ce dernier¹⁰¹ peut le conduire à le punir. Ces équilibres revêtent donc une stabilité théorique évidente, rendant la pratique du don incertaine par le simple fait que l'employeur réciproque craint d'être assimilé à un coopérateur stratégique cherchant à maximiser son gain.

L'incorporation des considérations de réciprocité des joueurs dans la détermination des salaires permet de prévoir les comportements des agents suivant leur type et d'anticiper l'efficacité à attendre de telles pratiques au sein de la relation d'emploi.

Les conclusions du modèle théorique confirment l'importance du sentiment de bienveillance réciproque dans l'efficacité de ces pratiques de don en apportant un fondement stable à l'observation d'une corrélation positive entre salaire et effort. Elles amènent cependant à relativiser l'idée de Fehr et al. (1996) selon laquelle les équilibres coopératifs ne seraient attribuables qu'à l'expression partagée d'un tel sentiment entre les deux joueurs. Il apparaît en effet que l'appariement entre un employeur stratégique et un employé réciproque peut aboutir à l'issue du jeu la plus enviable socialement dès lors que l'employé anticipe une intention bienveillante.

Le comportement d'équilibre de l'employé se révèle être unique car directement conditionné par la valeur de sa motivation intrinsèque pour la réciprocité. Le lien positif entre salaire et effort est justifié dès lors que la relation implique un employé suffisamment réciproque.

¹⁰¹ On rappelle que l'anticipation de l'employé sur la bienveillance de l'employeur s'écrit :

$$\lambda_{LEL}^{\bar{w}} = \frac{1}{2} [\bar{w} - \underline{w} - \rho'' C(\bar{e})].$$

L'employeur égoïste qui s'attend avec certitude à la maximisation de l'effort

après l'offre \bar{w} est vu comme faiblement bienveillant, à plus forte raison si le coût de l'effort est important.

Celui-ci adopte une stratégie rationnelle de récompense coûteuse suite à une offre de salaire élevée, assurant l'efficacité de la pratique du don. Cette relation positive entre salaire et effort semble cependant atteindre un maximum si l'on ne tient compte que des considérations de réciprocité dans les processus de décision des agents.

De plus, une source possible d'inefficacité de l'échange de cadeaux réside dans la position informationnelle de l'employeur qui doit décider ou non de pratiquer le don sur la base d'une simple anticipation sur les croyances de l'employé. Pour cette raison, deux joueurs identifiés comme réciproques peuvent aboutir à l'équilibre de Nash représentant l'issue la moins enviable en termes de gains aussi bien pour l'employeur que pour l'employé. Cette conclusion théorique peut éventuellement mettre en doute l'idée que la bienveillance réciproque suffit à assurer la coopération mutuelle dans le cadre de contrats, tels que la relation d'emploi, induisant une incomplétude concernant les contributions de chacune des parties. Cet effet indésirable peut être d'autant plus crédible que les types des joueurs ont été considérés jusqu'à présent de connaissance commune. Or, le comportement coopératif de l'employé dépend fortement de sa motivation intrinsèque pour la réciprocité. Introduire un défaut d'observabilité de la nature des joueurs n'aurait pour conséquence que de rendre moins probable encore la pratique du don de la part de l'employeur réciproque.

Doit-on en conclure dès à présent que les observations de Fehr et al. (1996), compatibles avec une interprétation de la bienveillance réciproque comme mécanisme permettant d'assurer la coopération dans un tel jeu, ne proviendraient finalement pas d'agents réellement motivés pour la réciprocité (particulièrement pour le premier joueur) ? Doit-on considérer que la réciprocité obtient son pouvoir explicatif le plus large principalement dans sa dimension négative (punition vs récompense, Falk & Gächter, 1998) ? Et en ce qui concerne les comportements de don et de récompense coûteux, l'approche par la réciprocité gagne-t-elle donc à être complétée par d'autres considérations extra matérielles comme l'altruisme ?

La réponse se doit d'être prudente pour l'instant. En effet, les résultats théoriques montrent tout de même que la bienveillance réciproque joue un rôle du point de vue de l'employeur puisque dans le cas où $(\bar{w} - \underline{w} - C(\bar{e})) < 0$, son comportement non coopératif est atténué au profit d'une probabilité positive d'offre de don (Cf. Graphe 2-3). Par ailleurs, l'analyse des comportements de l'employeur et de l'employé a simplement permis d'aboutir à une expression des seuils en fonction des paramètres de la relation d'emploi. Est-ce que cette dernière induit une plage de stabilité plus importante pour la coopération entre deux joueurs réciproques ? En d'autres termes, la valeur des seuils à partir desquels le don ne constitue plus

la stratégie optimale du point de vue de l'employeur constitue-t-elle une réelle contrainte ou s'adresse-t-elle à des types « pathologiques » de joueurs réciproques ? Quelle est la sensibilité de tels seuils à une variation des paramètres caractérisant la relation d'emploi ?

Pour répondre à ces questions qui émergent soit directement des conclusions théoriques du modèle soit de la manière dont les considérations de réciprocité ont été introduites, le chapitre 3 propose une simulation. La simulation permet non seulement d'adapter la paramétrisation à celle adoptée par Fehr et al. (1996) dans leur expérimentation, mais également de savoir si les comportements qu'ils ont observés proviennent réellement d'un soucis de bienveillance réciproque exprimé par leurs sujets.

ANNEXES AU CHAPITRE 2

ANNEXE 2-1 : STABILITÉ DE L'ÉQUILIBRE DE NASH DANS LE CADRE D'UN JEU D'ÉCHANGE DE CADEAU EN TROIS ÉTAPES IMPLIQUANT DES JOUEURS ÉGOÏSTES.

Le jeu séquentiel d'échange de cadeaux peut incorporer une étape supplémentaire permettant à l'employeur de sanctionner ou de récompenser le choix d'effort effectué par l'employé.

Soit $r \in \{-1;0;1\}$, paramètre de décision de l'employeur. Si $r = 1$, l'employeur sanctionne l'employé et dégrade son gain matériel d'un montant f en supportant un coût $k(1)$ représentant la mise en place d'un contrôle de l'effort de l'employé. Si $r = -1$, il choisit au contraire de récompenser l'employé d'un montant f , à un coût $k(-1)$ tel que $k(-1) > k(1)$. Le coût de la sanction est inférieur à celui de la récompense car ce dernier inclut f en supplément du coût de la mise en place du contrôle. On suppose enfin que $k(0) = 0$, l'employeur ne contrôle pas et n'intervient pas sur les gains matériels de l'employé. Quels que soient les choix de salaire et d'effort, les fonctions de gain de l'employeur et de l'employé s'écrivent ainsi :

pour l'employeur : $\pi^E = e(q - w) - k(r)$, avec $e \in \{\underline{e}; \bar{e}\}$; $w \in \{\underline{w}; \bar{w}\}$ et $r \in \{-1;0;1\}$

pour l'employé : $\pi^L = w - C(e) - rf$, avec $e \in \{\underline{e}; \bar{e}\}$; $w \in \{\underline{w}; \bar{w}\}$ et $r \in \{-1;0;1\}$

Les stratégies de récompense et de sanction coûteuses de l'employeur ne sont pas crédibles. Elles n'ont donc aucun impact sur le choix d'effort de l'employé.

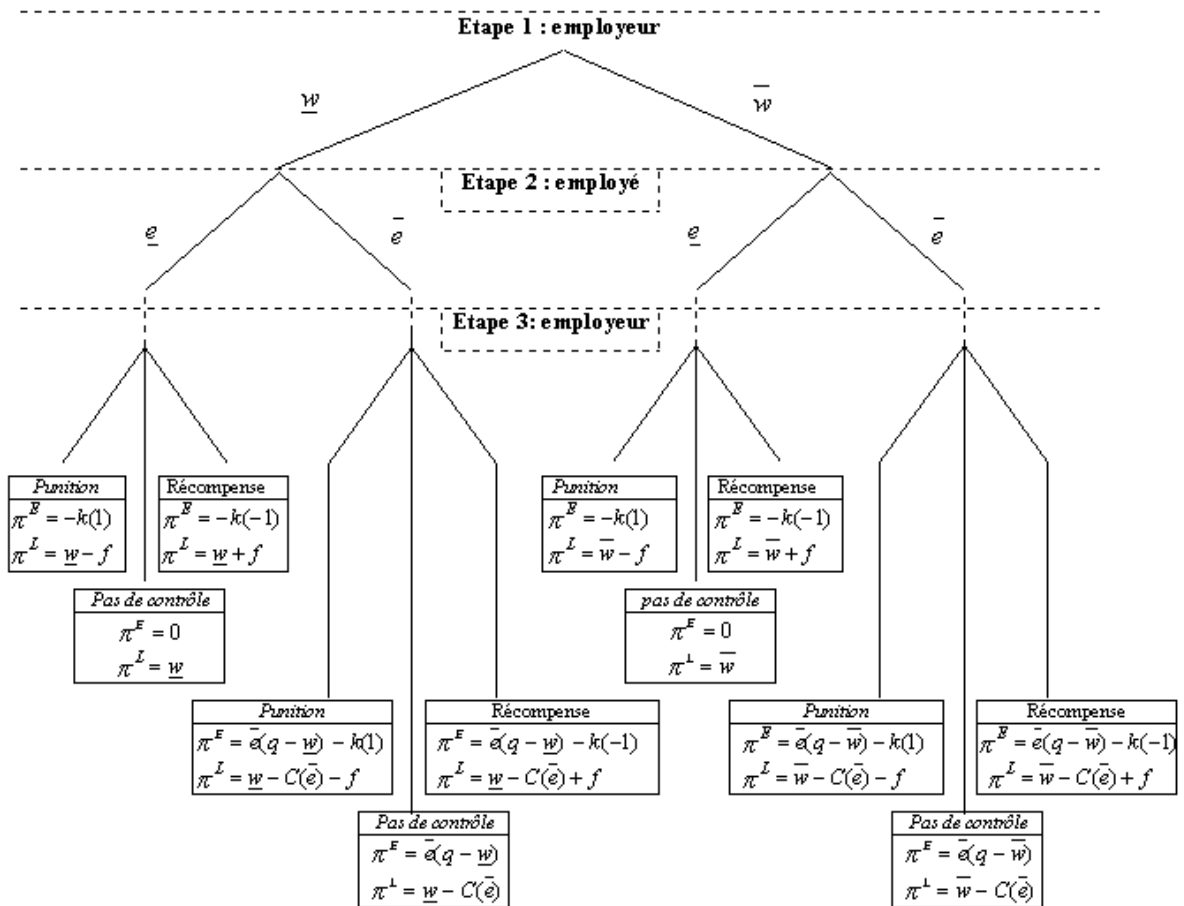
En effet, dès lors que ces deux actions sont coûteuses, la maximisation du gain matériel de l'employeur l'induit à choisir de ne pas pratiquer le contrôle, quelle que soit l'offre de salaire.

Par induction à rebours, l'employé anticipe l'absence du contrôle et sélectionne sa stratégie optimale : la dégradation de l'effort quelle que soit l'offre. Ainsi, au premier nœud du jeu, l'employeur décide de ne pas pratiquer le don (voir graphique suivant représentant le jeu sous forme extensive avec les paiements des joueurs).

Même en modifiant les fonctions de paiement de l'employeur de sorte qu'il puisse faire supporter l'intégralité du coût de contrôle à l'employé¹⁰², l'équilibre de Nash du jeu avec absence d'échange de don domine. En effet, en considérant que le coût de contrôle est donné, exogène, la fonction de paiement de l'employeur s'écrit :

$\pi^E = e(q - w) - k(r) + rf$, sachant que $k(1) = k(-1)$ (le coût du contrôle est le même, que l'employeur punisse ou récompense ensuite). On suppose en outre que l'employeur fait supporter à l'employé au moins l'intégralité du coût de contrôle s'il a dégradé son effort, ce qui implique $f \geq k(r)$. Si $f = k(r)$, l'employeur cherchant à maximiser son paiement de fin de jeu est indifférent entre "ne pas contrôler" et "contrôler et punir". Le coût associé au contrôle suivi de la récompense étant $2k(r)$, cette stratégie ne peut être sélectionnée. Si $f > k(r)$, il choisit systématiquement de contrôler et de punir, même si l'employé a fourni l'effort maximum. L'employé cherchant à maximiser son gain anticipe un tel comportement de la part de l'employeur et dégrade systématiquement son effort. Au premier nœud du jeu, l'employeur décide donc d'offrir le salaire faible. Ainsi, si les joueurs cherchent uniquement à maximiser leur gains de fin de jeu, un équilibre caractérisé par l'absence de don domine, que l'employeur puisse punir ou non. Cette troisième étape n'est qu'un *signal gratuit* et n'induit aucune conséquence comportementale.

¹⁰² Les fonctions de paiement précédentes, avec $k(-1) > k(1)$, impliquent que l'employeur récupère une partie des coûts de contrôle de l'effort lorsqu'il punit.



ANNEXE 2-2 : ECRITURE DE L'UTILITÉ DE L'EMPLOYEUR APPARIÉ AVEC UN EMPLOYÉ DONT LE TYPE EST AMBIGU

La fonction d'utilité de l'employeur apparié à un employé dont la motivation intrinsèque pour la réciprocité est ambiguë peut s'écrire de la même façon que lorsqu'il fait face à un employé réciproque. La différence réside dans la réaction de l'employé vis à vis de l'offre du don. Dans la situation précédente, l'employé suffisamment réciproque coopère à coup sûr alors que son type lui suggère à présent de maximiser son effort avec une probabilité non nulle ρ .

Au total, le type de l'employé est tel que $Y_L^{**} < Y_L < Y_L^*$ et que sa probabilité de coopérer

$$\text{face au don est } \rho(e = \bar{e} / w = \bar{w}) = \frac{\bar{w} - w}{C(\bar{e})} - \frac{2}{Y_L \bar{e}(q - \bar{w})}.$$

Au premier nœud du jeu, l'employeur, par son offre, peut procurer à l'employé soit \underline{w} , ou $\rho(\bar{w} - C(\bar{e})) + (1 - \rho)\bar{w}$. Du point de vue de l'employé, le gain équitable devient donc :

$$\pi_L^{eq} = \frac{1}{2}(\bar{w} + \underline{w} - \rho C(\bar{e})).$$

La bienveillance de l'employeur, pour chacune des propositions de salaire est définie ainsi :

$$K_{EL}^{\bar{w}} = \frac{1}{2}(\bar{w} - \underline{w} - \rho C(\bar{e})), \text{ s'il pratique le don; et}$$

$$K_{EL}^{\underline{w}} = -\frac{1}{2}(\bar{w} - \underline{w} - \rho C(\bar{e})) \text{ s'il opte pour l'hypothèse basse.}$$

Pour juger de la bienveillance de l'employé, l'employeur anticipe ses intentions face à l'offre qui lui est faite. Pour cela, il anticipe la croyance a priori de l'employé sur la proposition qui va effectivement lui parvenir. Cette croyance de deuxième ordre sur la probabilité de pratiquer le don est notée ε'' . Dans ce cas, l'employeur pense que l'employé lui procurera $\varepsilon''[\rho(\bar{e}(q - \bar{w})) + (1 - \rho)0] + (1 - \varepsilon'')0$ en choisissant sa stratégie d'équilibre. Si l'employé maximisait systématiquement son effort quelle que soit l'offre, le gain de l'employeur serait $\varepsilon''\bar{e}(q - \bar{w}) + (1 - \varepsilon'')\bar{e}(q - \underline{w})$, suivant la stratégie sélectionnée par ce dernier. Il serait de $\varepsilon''0 + (1 - \varepsilon'')0 = 0$ dans le cas contraire. Ainsi, la croyance de l'employeur sur la bienveillance de l'employé adoptant sa stratégie d'équilibre ($\lambda_{ELE}^{\{(w, e); (\bar{w}, p(e=\bar{e})=\rho)\}}$) peut être définie comme la différence entre son anticipation de gain et le gain équitable qu'il est en droit de revendiquer sachant que le type de l'employé est ambigu. On obtient donc :

$$\lambda_{ELE}^{\{(w, e); (\bar{w}, p(e=\bar{e})=\rho)\}} = \varepsilon''\bar{e}(q - \bar{w})(\rho - \frac{1}{2}) - \frac{1}{2}(1 - \varepsilon'')\bar{e}(q - \underline{w})$$

Ainsi, l'utilité de l'employeur pour chaque offre de salaire qu'il peut faire s'écrit :

S'il propose le don (salaire élevé) :

$$U_E^{\bar{w}} = \pi_E(\bar{w}, p(e = \bar{e}) = \rho) + Y_E K_{EL}^{\bar{w}} \lambda_{ELE}^{\{(w, e); (\bar{w}, p(e=\bar{e})=\rho)\}}, \text{ c'est-à-dire :}$$

$$U_E^{\bar{w}} = \rho\bar{e}(q - \bar{w}) + \frac{1}{2}Y_E(\bar{w} - \underline{w} - \rho C(\bar{e})) \left[\varepsilon''\bar{e}(q - \bar{w})(\rho - \frac{1}{2}) - \frac{1}{2}(1 - \varepsilon'')\bar{e}(q - \underline{w}) \right]$$

S'il propose le salaire faible :

$$U_E^{\underline{w}} = \pi_E(\underline{w}, e) + Y_E K_{EL}^{\underline{w}} \lambda_{ELE}^{\{(w, e); (\bar{w}, p(e=\bar{e})=\rho)\}}, \text{ c'est-à-dire :}$$

$$U_E^{\underline{w}} = -\frac{1}{2}Y_E(\bar{w} - \underline{w} - \rho C(\bar{e})) \left[\varepsilon''\bar{e}(q - \bar{w})(\rho - \frac{1}{2}) - \frac{1}{2}(1 - \varepsilon'')\bar{e}(q - \underline{w}) \right]$$

ANNEXE 2-3 : COMPORTEMENT D'ÉQUILIBRE DE L'EMPLOYEUR APPARIÉ AVEC UN EMPLOYÉ MODÉRÉMENT MOTIVÉ PAR LA RÉCIPROCITÉ

L'employeur propose \bar{w} si $U_E^{\bar{w}} > U_E^{\underline{w}}$, c'est-à-dire si :

$$Y_E(\bar{w} - \underline{w} - \rho C(\bar{e}))A(\rho, \varepsilon) > -\rho \bar{e}(q - \bar{w}), \quad (1)$$

$$\text{avec } A(\rho, \varepsilon) = \varepsilon'' \bar{e}(q - \bar{w})\left(\rho - \frac{1}{2}\right) - \frac{1}{2}(1 - \varepsilon'')\bar{e}(q - \underline{w})$$

L'examen des conditions d'une proposition de salaire élevé de la part de l'employeur est plus complexe dans ce cas car ε et ρ influencent directement le signe de $A(\cdot)$. Il existe un seuil minimum positif pour ε et le signe change si ρ atteint 0.5. Tous les cas sont envisagés :

Lorsque la probabilité que l'employé choisisse le niveau d'effort élevé est au moins égale à 0.5 : $\rho > 0.5$

pour que $A(\cdot)$ soit positif, le seuil minimum de probabilité de l'employeur doit être égal à

$$\mathcal{E}_0 = \frac{q - \underline{w}}{2\rho(q - \underline{w}) + (\bar{w} - \underline{w})}.$$

En utilisant la même procédure que pour le comportement d'équilibre de l'employeur apparié à un employé réciproque, les conditions pour que celui-ci propose le salaire élevé (faible) s'écrivent :

$$- \quad \text{Si } (\bar{w} - \underline{w} - \rho C(\bar{e})) > 0$$

Si l'employeur décide d'offrir le salaire élevé, la condition (1) doit être remplie pour $\varepsilon''=1$. Sa motivation pour la réciprocité doit être telle que :

$$Y_E > \frac{-\rho}{(\bar{w} - \underline{w} - \rho C(\bar{e}))(\rho - 0.5)} = Y_E^*, \text{ ce qui n'est pas contraignant puisque } Y_E^* < 0.$$

Par contre, s'il propose le salaire faible, la condition (1) ne doit pas être remplie pour $\varepsilon''=0$, ce qui implique :

$$Y_E > \frac{2\rho(q - \bar{w})}{(\bar{w} - \underline{w} - \rho C(\bar{e}))(q - \underline{w})} = Y_E^{**} > 0. \text{ Si sa préférence intrinsèque pour la réciprocité se}$$

situe en-dessous de Y_E^{**} , il propose le salaire élevé car il anticipe une réaction bienveillante de la part de l'employé. Au-delà, sa pratique du don est compromise car il s'attend à ce qu'une offre de salaire élevé soit interprétée comme un comportement stratégique de sa part.

- Si $(\bar{w} - \underline{w} - \rho C(\bar{e})) < 0$,

Alors il propose le salaire élevé si sa motivation intrinsèque pour la réciprocité est telle que ($\epsilon''=1$) :

$$Y_E < \frac{-\rho}{(\bar{w} - \underline{w} - \rho C(\bar{e}))(\rho - 0.5)} = Y_E^* > 0,$$

et le salaire faible si ($\epsilon''=0$) :

$$Y_E < \frac{2\rho(q - \bar{w})}{(\bar{w} - \underline{w} - \rho C(\bar{e}))(q - \underline{w})} = Y_E^{**} < 0, \text{ ce qui est impossible.}$$

En conséquence, l'employeur proposera à coup sûr le salaire élevé à un employé dont la probabilité de coopération est supérieure à 0.5 seulement si sa motivation intrinsèque pour la réciprocité est inférieure au seuil Y_E^* . D'autre part, si sa motivation pour la réciprocité dépasse ce seuil, il proposera le salaire élevé avec la probabilité ϵ' définie par :

$$\epsilon' = \frac{-2(q - \bar{w}) + Y_E(q - \underline{w})(\bar{w} - \underline{w} - C(\bar{e}))}{Y_E(\bar{w} - \underline{w} - \rho C(\bar{e}))(2\rho(q - \bar{w}) + \bar{w} - \underline{w})}. \text{ Il s'agit d'une écriture plus générale de la}$$

probabilité ϵ déjà exprimée puisque $\epsilon'(\rho=1)=\epsilon$ (c'est-à-dire la probabilité de proposer le salaire élevé quand il est apparié avec un employé réciproque).

Lorsque la probabilité que l'employé choisisse le niveau d'effort élevé est inférieure à 0.5 : $\rho < 0.5$

Dans ce cas, $A(.)$ est toujours négatif et il n'existe pas de minimum pour ϵ . On considère les deux cas précédents pour déterminer le comportement d'équilibre de l'employeur.

- Si $(\bar{w} - \underline{w} - \rho C(\bar{e})) > 0$,

L'employeur propose \bar{w} (condition (1) sachant $\epsilon''=1$) si :

$$Y_E < \frac{\rho}{(\bar{w} - \underline{w} - \rho C(\bar{e}))(\rho - 0.5)} = Y_E^* > 0,$$

et le salaire faible (\underline{w}) si : $Y_E > \frac{2\rho(q - \bar{w})}{(\bar{w} - \underline{w} - \rho C(\bar{e}))(q - \underline{w})} = Y_E^{**} > 0$

- Si $(\bar{w} - \underline{w} - \rho C(\bar{e})) < 0$,

L'offre de salaire élevé implique une motivation minimale pour la réciprocité telle que :

$Y_E > \frac{\rho}{(\bar{w} - \underline{w} - \rho C(\bar{e}))(\rho - 0.5)} = Y_E^*$, ce qui n'est pas contraignant puisque $Y_E^* < 0$,

et la proposition de salaire faible implique : $Y_E < \frac{2\rho(q - \bar{w})}{(\bar{w} - \underline{w} - \rho C(\bar{e}))(q - \underline{w})} = Y_E^{**} < 0$, ce qui est impossible.

En résumé, le comportement d'équilibre de l'employeur apparié avec un employé modérément réciproque est tel que :

		seuil de motivation pour la réciprocité de l'employeur	
Employé modérément réciproque		0	Y_E^* Y_E^{**}
probabilité de coopération : $\rho < 0.5$	$(\bar{w} - \underline{w} - \rho C(\bar{e})) > 0$	\bar{w}	\underline{w}
	$(\bar{w} - \underline{w} - \rho C(\bar{e})) < 0$	\bar{w}	
Probabilité de coopération : $\rho > 0.5$	$(\bar{w} - \underline{w} - \rho C(\bar{e})) > 0$	\bar{w}	\underline{w}
	$(\bar{w} - \underline{w} - \rho C(\bar{e})) < 0$	\bar{w}	
			$P(w = \bar{w}) = \varepsilon$

Les parties grisées correspondent aux valeurs du bénéfice lié à l'offre de salaire élevé écartées par hypothèse dans l'analyse. Aussi, le comportement d'offre de l'employeur en fonction de son type, lorsqu'il est apparié avec un employé modérément réciproque, n'est pas sensible à la probabilité de coopération de ce dernier. C'est uniquement le type de l'employé qui détermine l'offre de l'employeur.