

MÉGADONNÉES ET EXTERNALITÉS INFORMATIONNELLES*

Claude FLUET
Université Laval
CRÉFiR et CRED
claude.fluet@fsa.ulaval.ca

RÉSUMÉ – Une littérature récente soulève la question du degré souhaitable de visibilité ou publicité des comportements individuels. Dans ces analyses, une plus grande visibilité renforce les incitations réputationnelles, ce qui peut être socialement utile. En contrepartie, les comportements qui en résultent sont moins informatifs quant aux caractéristiques des individus, ce qui peut être indésirable. Je réexamine cette question dans un contexte où plus de visibilité découle d'un plus grand nombre de sources d'information sur les agissements des individus, dans de multiples interactions contractuelles. Contrairement à la littérature récente, plus de visibilité a alors pour effet de diluer ou même d'éliminer complètement les incitations réputationnelles. En contrepartie, cela améliore l'information sur les caractéristiques des individus, avec à la limite la disparition des asymétries d'information.

ABSTRACT – A recent literature raises the question of the desirable degree of visibility or publicity of individual behaviors. In these analyzes, greater visibility strengthens reputation incentives, which can be socially useful. In return, the resulting behaviors are less informative about the characteristics of individuals, which can be undesirable. I re-examine this question in a context where more visibility stems from a greater number of sources of information on the actions of individuals, in multiple contractual interactions. Contrary to recent literature, more visibility then has the effect of diluting or even completely eliminating reputational incentives. In return, this improves information on the characteristics of individuals, with the ultimate disappearance of information asymmetries.

INTRODUCTION

Les nouvelles technologies de collecte, d'exploration et de circulation de données sont de plus en plus utilisées pour classer et évaluer les individus. L'analyse

*Cet article s'inspire d'une communication présentée en séance plénière au 59e Congrès annuel de la Société Canadienne de Science Économique, Québec, 8-10 mai 2019, sous le titre *Mégadonnées, discrimination statistique et incitations*. Je remercie un rapporteur pour ses commentaires sur une première version de ce texte.

prédictive à partir de mégadonnées est employée pour les décisions de recrutement, l'estimation du risque de défaut dans l'octroi de crédits, l'évaluation de la fiabilité des locataires, le ciblage de la publicité et l'optimisation de l'offre de produits, la détection de la fraude à l'assurance, l'estimation du risque de récidive, etc. Ces pratiques ont fait l'objet de très nombreuses critiques et mises en gardes, par exemple en ce qui concerne l'opacité des procédures (Federal Trade Commission, 2016), le profilage des consommateurs à des fins de manipulation et discrimination de prix (Acquisti *et al.*, 2016) ou les effets indésirables de la discrimination statistique. En particulier, l'attribution à un individu des qualités de ceux qui lui sont statistiquement proches aurait pour conséquence de pérenniser, voire d'accroître les inégalités aux dépens de groupes historiquement désavantagés. Cette préoccupation a été dominante dans la littérature juridique (Barocas et Selbst, 2016, entre autres). Elle tend aussi à être maintenant prise en compte dans le développement des algorithmes prédictifs.¹

Une autre préoccupation, non directement reliée aux techniques de prédiction par corrélations, concerne le degré optimal de protection de la vie privée. Dans deux articles souvent cités, Posner (1977, 1981) argumente que le maintien du caractère confidentiel de certaines caractéristiques individuelles ou de la confidentialité des agissements passés d'un individu est source d'asymétries d'information, avec des effets indésirables dans les relations contractuelles ou sociales.² Sur ce plan, les nouvelles technologies ont manifestement le potentiel d'augmenter considérablement la visibilité des caractéristiques ou comportements individuels. Pour certains, cela représente un danger et aboutirait à une *scored society* ou *reputation nation*, pour reprendre des titres d'articles évocateurs (Citron et Pasquale, 2014; Strahilevitz, 2008). On peut évoquer aussi le *Système de crédit social* mis en place par le gouvernement chinois à partir de 2014 (voir Creemers, 2018).

Une littérature économique récente soulève la question du degré socialement optimal de transparence ou de visibilité. La particularité de cette littérature est d'aborder la question sous l'angle des motivations réputationnelles des individus. Réputationnel signifie ici que les sanctions ou bénéfices associés à certains comportements dépendent des croyances qui se formeront *ex post* relativement aux caractéristiques des individus étant donné leurs agissements passés, par opposition à des sanctions ou bénéfices prédéterminés comme dans un régime pénal ou un contrat de performance classique. Dans ces analyses, une plus grande visibilité des comportements renforce les incitations réputationnelles. Celles-ci peuvent être

1. La revue de littérature multidisciplinaire de Romei et Ruggieri (2014) sur la discrimination statistique contient plus de 500 références, dont des classiques chez les économistes (Phelps *et al.*, 1972; Arrow, 1973; Fang et Moro, 2011), mais aussi des textes de juristes et de chercheurs en analyse prédictive. Les travaux récents portent sur la mesure des discriminations statistiques, relativement aux critères faisant l'objet d'une protection légale, et l'arbitrage entre discrimination et efficacité prédictive (par exemple d'Alessandro *et al.*, 2017; Žliobaitė, 2017)

2. Dans des relations avec aléa moral, il est bien connu qu'une meilleure information sur les comportements atténue le conflit entre partage de risque et incitations (Hölmstrom, 1979) ou diminue les distorsions dues à l'extraction de rente (Demougin et Fluet, 2001). De même, une meilleure catégorisation des individus diminue les inefficacités dues à l'antisélection (Crocker et Snow, 1986).

socialement utiles, par exemple lorsqu'il s'agit d'inciter les individus à contribuer à un bien public ou à s'abstenir d'actions avec des externalités négatives (Bénabou et Tirole, 2006, entre autres). Toutefois, il y a généralement un coût. Daughety et Reinganum (2010) analysent l'arbitrage entre l'inefficacité des actions d'un point de vue individuel, due à l'exacerbation des incitations réputationnelles, et les externalités positives qui leur sont associées. Ainsi, lorsque les actions sont une information purement privée, les contributions à bien public seraient insuffisantes; en revanche, lorsque les actions sont parfaitement publicisées, les contributions pourraient être excessives. Mazyaki et van der Weele (2019) montrent également que les incitations réputationnelles sont difficiles à maîtriser et peuvent conduire à des comportements excessifs, du fait notamment de la multiplicité des équilibres possibles.

Une autre considération est que les incitations réputationnelles se traduisent par des comportements moins informatifs sur les caractéristiques des individus. Ali et Bénabou (2016) analysent le problème d'un régulateur cherchant à induire des comportements socialement désirables, mais qui est imparfaitement informé des préférences véritables des agents, de sorte qu'il n'est pas certain de la politique à adopter. Dans la situation considérée, les motivations réputationnelles sont utiles à des fins incitatives, mais les équilibres qui en découlent nuisent à l'extraction d'information sur les préférences.³ Le régulateur choisit le niveau de publicité des comportements en arbitrant entre les incitations réputationnelles et l'extraction d'information. Chez Jann et Schottmüller (2020), les individus travestissent les opinions qu'ils expriment parce qu'ils craignent de faire l'objet de discrimination statistique, ce qui nuit à l'agrégation d'information sur les caractéristiques des individus. Chez Deffains et Fluet (2020), un régulateur cherche à discipliner les comportements par la qualification des infractions (c'est-à-dire la définition de ce qui constitue une infraction) et par une politique de détection et de sanctions. Les incitations réputationnelles diminuent les coûts de régulation, mais le régulateur doit arbitrer entre le recours à ces incitations et la qualité de l'information générée par le mécanisme disciplinaire relativement aux caractéristiques des agents. Dans leur analyse, cette information a une valeur sociale dans les interactions futures, contractuelles ou autres, impliquant les agents.

Dans les travaux que je viens de décrire, une plus grande publicité des comportements conduit à des effets incitatifs plus forts, mais aux dépens de l'information qui peut être extraite de l'observation de ces mêmes comportements. Dans ce qui suit, j'analyse une situation où la publicisation des comportements se produit dans un cadre décentralisé. Un individu est impliqué dans de multiples relations de type principal-agent. Dans chacune de ces interactions, le comportement ou la performance de l'individu pourraient demeurer une information privée, en ce sens qu'elle n'est connue que des participants à la relation, ou au contraire l'infor-

3. Un exemple de cet effet est l'excès de conformisme analysé par Bernheim (1994) et Morris (2001). Néanmoins, la diminution de la valeur informationnelle des comportements ne passe pas nécessairement par le conformisme; voir Fox et Van Weelden (2012).

mation pourrait être rendue publique, ce qui signifie qu'elle est communiquée à d'autres principaux. Dans ce contexte, plus de visibilité peut résulter de l'augmentation du nombre de sources d'observation, c'est-à-dire du nombre d'interactions contractuelles communiquant un score de performance sur l'individu, ce qui rejoint l'idée de *scored society* mentionnée ci-dessus. Les individus ont donc à gérer les conséquences réputationnelles de leur comportement sur de multiples relations contractuelles. Je suppose qu'ils ont une capacité limitée de coordination de leurs actions à cet égard. Plus de visibilité a alors l'effet contraire de celui qui est décrit au paragraphe précédent. Passé un certain seuil, une plus grande visibilité diminue la force des incitations réputationnelles, jusqu'à potentiellement les réduire à zéro, et améliore la qualité de l'information extraite de l'observation des comportements. La conclusion de cette analyse est donc que plus de visibilité ou de transparence, par l'agrégation de sources d'information multiples sur les individus, réduit ou élimine les asymétries d'information mais a pour effet de diluer les incitations. Il y a par conséquent un seuil optimal de visibilité, comme dans la littérature évoquée ci-dessus, mais pour des raisons diamétralement opposées.

La section 1 développe un modèle simple de relation contractuelle de type principal-agent avec aléa moral et antisélection, d'abord sans incitations réputationnelles et en introduisant ensuite de telles incitations dans le cadre de cette seule relation. Ce modèle me permet d'analyser les externalités informationnelles entre un secteur communiquant de l'information sur les individus, sous la forme de scores de performance, et un secteur valorisant cette information dans le cadre de relations contractuelles futures. La section 2 introduit la possibilité de multiples sources d'information, sous la forme de plusieurs relations principal-agent connectées à des échanges d'information, et dégage les principaux résultats. La section 3 conclut.

1. EXTERNALITÉS INFORMATIONNELLES

Considérons une relation entre un principal et un agent, par exemple une relation employeur-employé, locateur-locataire, prêteur-emprunteur, etc. De façon schématique, le surplus de la relation est

$$S = y + (h - g)e, \quad e = 0, 1,$$

où e est le comportement de l'agent, h un gain pour le principal (ou un dommage évité) et g le coût d'opportunité pour l'agent d'adopter le comportement $e = 1$. Le premier terme pourrait se décomposer en $y = y_A + y_P$, le surplus de la relation étant la somme du gain pour l'agent, $y_A - ge$, et du gain pour le principal, $y_P + he$. Je suppose que $g < h$, de sorte que le surplus est maximisé si l'agent choisit l'action $e = 1$. Le problème, quel que soit le partage du terme y , est que l'agent supporte le coût de l'action (diligence au travail, soin apporté à un logement locatif, réclamation honnête présentée à un assureur, etc.) sans en tirer tout le bénéfice.

Il y a deux types d'agent, le bon agent dénoté b et le mauvais agent dénoté m . La probabilité d'avoir affaire à un bon agent est λ , celle d'avoir affaire à un mauvais agent est $1 - \lambda$. Pour faire simple, je suppose que le bon agent n'a pas de coût d'opportunité : il adopte toujours le comportement qui maximise le surplus de la relation.⁴ En revanche, le mauvais agent a un coût g qui dépend des circonstances aléatoires auxquelles il fera face, définies par la fonction de répartition $F(g)$ avec densité $f(g)$ sur le support $[0, h]$. Le type de l'agent et le coût d'effort s'il s'agit d'un mauvais agent sont inobservables par le principal. Dans ce qui suit, j'analyse la gouvernance de la relation, en élargissant graduellement le cadre de l'analyse.

Gouvernance. Le principal vérifie les comportements avec une probabilité p moyennant une dépense $c(p)$. Il peut détecter si un agent a causé des dommages, s'il a fait une déclaration frauduleuse, etc. Le coût d'audit satisfait $c(p) = c'(p) = 0$ pour $p \leq \bar{p}$ et $c'(p) > 0$ sinon, avec $\bar{p} \in (0, 1)$. Un minimum d'information est donc disponible sans frais, mais une plus grande surveillance nécessiterait une dépense.

Si un manquement est détecté, le principal impose à l'agent la pénalité monétaire s . Je fais l'hypothèse que $\bar{p}s < h$. Par conséquent, une stratégie d'audit passive, où le principal n'investit pas en détection, ne suffit pas à bien inciter le mauvais agent dans toutes les circonstances auxquelles il pourrait faire face. L'audit actif étant coûteux, il serait dans l'intérêt de la relation que la sanction soit la plus élevée possible. J'interprète s comme représentant ce maximum, considéré comme exogène. Par exemple, dans une relation locateur-locataire, la sanction est la perte de la caution versée par le locataire et il y a une limite aux cautions exigibles.

Comme le bon agent a toujours le comportement adéquat, il n'est jamais sanctionné. Quant au mauvais agent, il choisit de faire l'effort ou non en fonction de son coût d'effort et du risque de sanction. Il a donc une stratégie de seuil : il fait l'effort lorsque $g \leq ps$ et ne le fait pas lorsque $g > ps$. Soit g_c le seuil de coût au-delà duquel l'agent adopte un comportement inefficace, donc $g_c = ps$. La probabilité d'audit détermine ainsi la probabilité $F(g_c)$ que le mauvais agent se comporte de manière efficace du point de vue de la relation.

4. Une autre formulation serait que le bon agent a un comportement normatif internalisant les effets de ses actions sur le surplus de la relation. Cette interprétation est standard dans la littérature sur les normes sociales et les motivations intrinsèques, par exemple Bénabou et Tirole (2003, 2006), Besley et Ghatak (2005), Akerlof et Kranton (2005), Ramalingam et Rauh (2010). Pour des études expérimentales montrant le rôle de ces motivations, voir Charness et Rabin (2002), Charness *et al.* (2016) ou Deffains *et al.* (2019).

Le surplus attendu de la relation, net du coût d'audit et compte tenu de la proportion de bons et mauvais agents, est⁵

$$S = y + \lambda h + (1 - \lambda) \int_0^{g_c} (h - g) f(g) dg - c(p). \quad (1)$$

La gouvernance optimale de la relation consiste à choisir la probabilité d'audit de façon à maximiser le surplus attendu. On a

$$\left. \frac{\partial S}{\partial p} \right|_{p \leq \bar{p}} = (1 - \lambda)(h - ps)f(ps)s > 0$$

et

$$\left. \frac{\partial S}{\partial p} \right|_{p = h/s} = -c'(h/s) < 0.$$

La solution est donc une probabilité $p \in (\bar{p}, h/s)$ satisfaisant la condition de premier ordre

$$c'(p) = (1 - \lambda)(h - ps)f(ps)s. \quad (2)$$

L'audit optimal égalise le coût marginal de détection au gain marginal découlant des effets incitatifs de l'augmentation de l'espérance de sanction.⁶ Soit p_0 la probabilité optimale et S_0 le surplus qui lui est associé. Comme $p_0 s < h$, le comportement du mauvais agent sera donc parfois inefficace, selon les circonstances auxquelles il est confronté. Comparée à l'optimum de premier rang, la perte de surplus résulte de cette inefficacité et du coût du mécanisme de gouvernance.

Incitations réputationnelles. J'intègre maintenant ce qui précède dans un ensemble plus large. Premièrement, il y a un grand nombre de relations contractuelles, impliquant des agents et des principaux différents, mais toujours avec les mêmes caractéristiques et la même proportion λ de bons agents. J'appelle Système I l'ensemble de ces relations. Deuxièmement, les agents auront à une date ultérieure d'autres interactions ou relations contractuelles, dans le cadre de ce que je désigne par Système II.

En l'absence d'échanges d'information entre les systèmes, les équilibres dans le Système I correspondent à l'équilibre décrit ci-dessus. J'introduis cependant la possibilité que la performance d'un agent génère un signal, dénoté X , transmis aux cocontractants potentiels du Système II. Ce signal est binaire : la réalisation

5. La sanction s est un pur transfert, payé par l'agent et reçu par le principal. Elle n'affecte donc pas le surplus de la relation, sauf par le biais des effets incitatifs du risque de sanction.

6. Dans l'équation (2), $f(ps)s = dF(ps)/dp$ est la variation de la probabilité que le mauvais agent fasse l'effort quand la probabilité de détection augmente.

$X = M$ signifie que l'agent a été sanctionné dans le Système I ou, si l'on préfère, qu'il a été mal noté ou a fait l'objet d'un blâme ; la réalisation $X = N$ signifie que l'agent n'a pas été sanctionné. Autrement dit, M est un score dénotant la « mauvaise performance », alors que N est « neutre » sur ce plan. La notation de l'agent détermine la croyance μ des cocontractants du Système II quant au type de l'agent, μ désignant la probabilité a posteriori que l'agent soit du bon type.

Soit μ_M la croyance quand l'agent a été mal noté et μ_N la croyance quand il n'a pas été mal noté. Un agent du bon type n'est jamais sanctionné, donc jamais mal noté. Le mauvais agent n'est pas mal noté s'il ne subit pas d'audit ou s'il en subit un lorsqu'il a le comportement adéquat. Si g_c est le seuil décrivant le comportement des mauvais agents, la probabilité qu'un mauvais agent ne soit pas mal noté est égale à $1 - p + pF(g_c)$. En appliquant la règle de Bayes, les croyances satisfont donc

$$\mu_M = 0, \quad (3)$$

$$\mu_N = \frac{\lambda}{\lambda + (1 - \lambda)[1 - p + pF(g_c)]}. \quad (4)$$

L'équation (4) détermine μ_N pour un seuil g_c et une probabilité p quelconques. Les valeurs d'équilibre seront dérivées ci-dessous.

Un blâme révèle toujours parfaitement que l'agent est du mauvais type. L'absence de blâme, cependant, peut être plus ou moins favorable à l'agent, puisque cela dépend de la probabilité d'audit et du comportement des mauvais agents. L'équation (4) définit une fonction $\mu_N(p, g_c)$ croissante en p et décroissante en g_c tant que $g_c < h$. Dans le cas limite où $g_c = h$, les mauvais agents ne sont jamais sanctionnés et on a $\mu_N(p, h) = \lambda$ quelle que soit la probabilité d'audit. Les mauvais agents se conduisent alors exactement comme les bons agents, de sorte que l'absence de blâme ne fournit aucune information sur leur type.⁷

L'utilité que retire un agent de ses interactions dans le Système II dépend des croyances à son égard, ce que je représente par la fonction $v(\mu)$. Je discuterai plus loin des propriétés de cette fonction. Pour l'instant, il suffit qu'elle soit croissante : les gains d'un agent dans ses relations ou transactions futures dépendent de la perception que l'on a de son type et donc la note qu'il a reçue. Le bon agent se comporte toujours de manière moins opportuniste, de sorte que $v(\mu)$ est une fonction croissante. Le notation des agents modifiera les actions dans le Système I, à la fois pour les mauvais agents dans leur choix de comportement et pour les principaux dans leur stratégie d'audit.

7. Le score $X = M$ est alors hors équilibre et μ_M ne peut être obtenu par la règle de Bayes. La valeur $\mu_M = 0$ est cependant compatible avec le critère D1 de Cho et Kreps (1987). Elle peut aussi s'obtenir en laissant tendre g_c vers h .

Considérons d'abord la stratégie des mauvais agents face à une probabilité d'audit quelconque de la part des principaux. Pour un coût d'effort g , le gain total à faire l'effort est maintenant

$$-g + v(\mu_N).$$

Lorsqu'il fait l'effort, l'agent supporte le coût g et est assuré de ne pas être mal noté, ce qui lui procure l'utilité réputationnelle $v(\mu_N)$. Lorsqu'il ne fait pas l'effort, le gain attendu est

$$p[v(0) - s] + (1 - p)v(\mu_N).$$

Avec la probabilité p , le comportement de l'agent est vérifié, auquel cas la sanction s lui est imposée et son utilité réputationnelle est $v(\mu_M) = v(0)$. Avec la probabilité $1 - p$, son comportement n'est pas vérifié et l'agent n'est pas mal noté. En comparant les deux gains, le mauvais agent choisit de faire l'effort si

$$g \leq p[s + v(\mu_N) - v(0)]. \quad (5)$$

Je dénoterai par $\Delta \equiv v(\mu_N) - v(0)$ la pénalité réputationnelle associée à une mauvaise note, comparativement à l'absence de blâme. Du point de vue des incitations, la pénalité réputationnelle s'ajoute à la sanction matérielle s imposée par le principal.

Pour une stratégie d'audit donnée, un équilibre réputationnel dans les relations contractuelles du Système I est défini par une stratégie g_c et une pénalité réputationnelle Δ satisfaisant le système d'équations :

$$g_c = \min[p(s + \Delta), h], \quad (6)$$

$$\Delta = v[\mu_N(p, g_c)] - v(0). \quad (7)$$

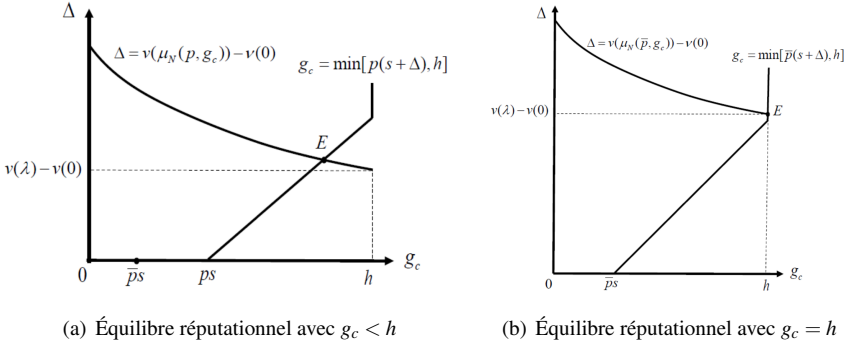
L'équation (6) est la fonction de seuil découlant de la condition (5), mais en intégrant la possibilité que la pénalité réputationnelle soit suffisamment élevée pour que les mauvais agents adoptent le comportement efficace. L'équation (7) donne la pénalité réputationnelle compatible avec le comportement des mauvais agents. Une solution de ce système d'équations est un Équilibre Bayesian parfait où le comportement de chaque agent dépend de celui de tous les autres, à travers l'effet du comportement moyen sur la pénalité réputationnelle.

L'équilibre est unique.⁸ Dans le graphique 1a, l'équilibre réputationnel correspond au point E et se caractérise par $g_c < h$ et $\Delta > v(\lambda) - v(0)$. Une augmentation de la probabilité d'audit p déplacerait la fonction de seuil vers la droite et entraîne-

8. L'unicité découle du fait que les comportements des agents sont des substituts stratégiques. Si les autres se comportent mieux, c'est-à-dire quand g_c est plus grand, la pénalité réputationnelle Δ diminue, ce qui est désincitatif.

rait une rotation de la courbe de pénalité dans le sens des aiguilles d’une montre ; voir l’équation (4). Les deux effets vont dans le sens d’une augmentation de g_c : lorsque la probabilité d’audit augmente, les mauvais agents sont plus fortement incités à adopter un comportement efficace.⁹

GRAPHIQUE 1
ÉQUILIBRE RÉPUTATIONNEL



Dans le graphique 1b, l’équilibre E correspond à une situation où $g_c = h$ et $\Delta = v(\lambda) - v(0)$. Dans le graphique, on a aussi supposé que cet équilibre est obtenu avec une stratégie d’audit passive $p = \bar{p}$. Comme les dépenses de détection sont nulles et que les mauvais agents ont le comportement efficace, le surplus des relations contractuelles est celui de l’optimum de premier rang. Autrement dit, la perte de surplus due à l’aléa moral est alors réduite à zéro. Dans la situation représentée, la probabilité d’audit à coût nul suffit pour inciter parfaitement les mauvais agents parce que les pénalités réputationnelles sont suffisamment fortes. Tel que représenté dans la figure, une probabilité d’audit légèrement inférieure à \bar{p} , ou une sanction légèrement inférieure à s , conduirait aussi au même résultat.

Je suppose que les principaux considèrent comme donnée la valeur réputationnelle des scores M et N . L’interprétation est que, dans les échanges d’information, les notations provenant de différentes relations principal-agent sont agrégées sur l’ensemble des relations. Par conséquent, chaque principal considère que sa stratégie d’audit n’a pas d’influence sur l’interprétation du signal qu’elle génère.¹⁰

À l’équilibre, selon les paramètres de la situation, les principaux adoptent soit une stratégie d’audit active menant à une situation comme celle du graphique 1a,

9. L’effet sur la pénalité réputationnelle est ambigu.

10. Dans Deffains et Fluet (2020), le principal est un régulateur qui prend en compte l’effet de la probabilité de détection sur le contenu du signal. Cela modifie l’arbitrage entre les coûts et les bénéfices de la vérification des comportements. On pourrait intégrer cet effet sans rien changer d’essentiel dans ce qui suit.

soit une stratégie passive menant à une situation comme celle du graphique 1b. L'équilibre est une solution en g_c , Δ et p des équations (6) et (7) combinées à la condition

$$c'(p) = (1 - \lambda)\{h - \min[p(s + \Delta), h]\}f(\min[p(s + \Delta), h])(s + \Delta). \quad (8)$$

Cette condition est de même nature que la condition (2). Chaque principal choisit la probabilité p de manière à égaliser le coût marginal d'audit au gain marginal en termes d'incitations. Dans ce calcul, la pénalité réputationnelle est prise comme donnée. Je dénote par p_1 la probabilité d'audit à l'équilibre et par S_1 le surplus attendu de la relation, obtenu en substituant les valeurs appropriées dans (1).

Proposition 1. *En présence d'effets réputationnels, (i) les relations contractuelles se caractérisent par $p_1 = \bar{p}$ et $g_c = h$ si $\bar{p}[s + v(\lambda) - v(0)] \geq h$ et par $p_1 > \bar{p}$ et $g_c < h$ si $\bar{p}[s + v(\lambda) - v(0)] < h$; (ii) dans les deux cas, le surplus S_1 de la relation est supérieur au surplus S_0 en l'absence d'effets réputationnels.*

Démonstration. L'énoncé (i) découle des équations (6), (7) et (8), sachant que $c'(\bar{p}) = 0$ et que $\Delta = v(\lambda) - v(0)$ lorsque $g_c = h$. Pour démontrer (ii), on utilise le fait que le surplus maximisé est

$$S^* \equiv \max_p y + \lambda h + (1 - \lambda) \int_0^{p\sigma} (h - g)f(g)dg - c(p)$$

où $\sigma = s + \Delta$. En l'absence d'effets réputationnels, on aurait $\sigma = s$. Par le théorème de l'enveloppe,

$$\frac{dS^*}{d\sigma} = \frac{\partial S^*}{\partial \sigma} = (1 - \lambda)(h - p\sigma)f(p\sigma)p > 0$$

où l'inégalité découle de (2) ou (8) pour le cas où $p\sigma < h$. Comme $\Delta \geq v(\lambda) - v(0) > 0$ dans tout équilibre, la conclusion suit. ■

On peut faire remarquer que l'augmentation du surplus due aux effets réputationnels provient soit de la diminution de la probabilité d'audit, donc des coûts de vérification, soit du comportement plus efficace des mauvais types (une augmentation de g_c), ou des deux.

Valeur de l'information. Le Système I communique un signal sur la performance des agents. Ce signal a une valeur *prédictive* pour les interactions impliquant les mêmes agents dans le Système II. En retour, le second système transmet une valorisation des signaux, ce qui détermine leurs conséquences *réputationnelles*, en fonction des effets potentiels dans le second système. J'analyse ici la relation entre valeur prédictive et valeur réputationnelle.

Pour simplifier, je suppose que les interactions ou relations contractuelles dans le Système II sont toutes de même nature et que chaque agent est impliqué dans une seule de ces interactions. Soit $q_t(a)$ le surplus d'une relation avec un agent

de type t , où t est m ou b , étant donné un vecteur a de variables de décisions décrivant la gouvernance de la relation. Je suppose que $q_b(a) > q_m(a)$ pour tout a : le surplus d'une relation avec un agent du bon type est toujours plus grand qu'avec un mauvais type. Lorsque le type est inobservable, mais qu'il y a une probabilité μ d'avoir affaire au bon type, le surplus attendu de la relation est

$$Q(\mu) \equiv \max_a \mu q_b(a) + (1 - \mu)q_m(a). \quad (9)$$

La fonction $Q(\mu)$ est croissante et convexe.¹¹

Je suppose que l'utilité réputationnelle des agents est une fonction croissante du surplus des relations dans lesquelles ils seront impliqués. Par exemple, $v(\mu) = \rho Q(\mu)$ où $\rho \in (0, 1]$ est la part des agents dans le surplus des relations contractuelles du Système II, la part des cocontractants étant alors égale à $1 - \rho$.

La probabilité a posteriori μ relative à un agent dépend du signal X transmis par le Système I et peut donc s'écrire comme une fonction $\mu(X)$. Ex ante, $\mu(X)$ est une variable aléatoire dont l'espérance est égale à λ , la probabilité a priori qu'un agent soit du bon type. En moyenne, sur l'ensemble des agents, la valeur des relations contractuelles dans le Système II est donc l'espérance de surplus :

$$\mathbb{E}[Q(\mu(X))] = P_M Q(\mu_M) + P_N Q(\mu_N), \quad (10)$$

où P_M est la probabilité de la note M et P_N la probabilité de la note N , avec évidemment $P_M = 1 - P_N$. La probabilité d'observer N est le dénominateur de l'expression dans l'équation (4),

$$P_N = \lambda + (1 - \lambda)[1 - p + pF(g_c)]. \quad (11)$$

Comme Q est convexe,

$$\mathbb{E}[Q(\mu(X))] \geq Q(\lambda) \quad (12)$$

avec inégalité stricte si Q est strictement convexe. Le terme à droite est le surplus des relations du Système II en l'absence de signaux. Le terme à gauche est le surplus moyen lorsque les principaux du Système II peuvent conditionner leurs décisions de gouvernance sur la performance des agents.¹²

11. La croissance découle directement du théorème de l'enveloppe, sachant que $q_b(a) > q_m(a)$ pour tout a . La convexité est un résultat classique en théorie de la décision ; voir Gollier (2001), chapitre 24. Par exemple, dans la relation principal-agent examinée précédemment (en l'absence d'effets réputationnels), la variable de gouvernance est p . En dénotant le surplus maximisé par $S(\lambda)$, on peut vérifier que $S'(\lambda) > 0$ et $S''(\lambda) > 0$.

12. Si un signal X' est plus informatif que X par rapport aux types des agents, la variable aléatoire $\mu(X')$ est un accroissement de risque relativement à $\mu(X)$; voir Ganuza et Penalva (2010). La convexité de Q implique alors $\mathbb{E}[Q(\mu(Y))] \geq \mathbb{E}[Q(\mu(X))]$. L'inégalité (12) en est un cas particulier.

La notation des agents dans le Système I augmente le surplus moyen des relations du Système II lorsque l'inégalité en (12) est stricte. Elle ne l'est pas si la solution du problème (9) conduit toujours à la même action, disons \hat{a} , de sorte que

$$\begin{aligned}\mathbb{E}[Q(\mu(X))] &= \mathbb{E}\{\mu(X)q_b(\hat{a}) + (1 - \mu(X))q_m(\hat{a})\} \\ &= \lambda q_b(\hat{a}) + (1 - \lambda)q_m(\hat{a}) \\ &= Q(\lambda).\end{aligned}$$

Je dirai que le signal a alors des effets purement redistributifs. Lorsque le signal influence les décisions de gouvernance, il a une valeur productive et l'inégalité en (12) est stricte.¹³

Proposition 2. *L'échange d'information entre les Systèmes I et II est mutuellement bénéfique. Le gain est strict pour les relations contractuelles du Système I; il est strict pour celles du Système II, sauf effets purement redistributifs de l'information ou lorsque le Système I atteint l'optimum de premier rang.*

Démonstration. Je complète l'argumentation dans le texte. Lorsque le Système I atteint l'optimum de premier rang, $P_N = 1$ et $\mu_N = \lambda$. Le surplus moyen des relations dans le Système II est alors

$$P_N Q(\mu_N) + (1 - P_N) Q(\mu_M) = Q(\lambda),$$

ce qui est identique au surplus en l'absence d'information provenant du Système I. ■

La proposition souligne que la valeur incitative des réputations ne va pas nécessairement de pair avec la valeur productive des prédictions. D'une part, la notation des agents a une utilité incitative même si les prédictions ont des effets purement redistributifs. D'autre part, lorsque les prédictions ont potentiellement une utilité productive, un signal prédictif ne peut être utile que s'il permet la différenciation des types. Les incitations réputationnelles reposent sur la menace d'une mauvaise réputation et il n'est pas nécessaire que cette menace soit mise à exécution. Elle ne l'est pas ici lorsque le Système I atteint l'optimum de premier rang. Inversement, la différenciation des types exige que les menaces soient mises à exécution avec une certaine probabilité. Les intérêts des deux systèmes ne sont donc pas parfaitement alignés.

On peut faire remarquer que, dans la situation considérée, la production d'information prédictive n'est pas optimisée. Les échanges d'information constituent des externalités, les bénéfiques pour les tiers n'étant pas pris en compte dans les stratégies des principaux du Système I.¹⁴ De même, le Système I profite de l'exis-

13. Cela n'empêche pas qu'il puisse y avoir une demande pour des informations avec des effets purement redistributifs. Voir Hirshleifer (1971).

14. Dans Deffains et Fluet (2020), le régulateur prend en compte la valeur productive de l'information transmise à des tiers dans ses décisions de régulation. Cela peut l'amener, par exemple, à

tence du Système II. Dans ce qui suit, d'autres sources d'externalités informationnelles sont introduites.

2. SIGNAUX MULTIPLES

J'élargis maintenant l'univers des relations contractuelles. Le Système I est composé de K secteurs, un nombre qui pourrait être très grand. Chacun de ces secteurs est constitué de relations principal-agent avec les mêmes caractéristiques que celles de la section précédente. Un agent est impliqué simultanément dans une relation contractuelle de chaque secteur. Dans chacune de ces relations, le mauvais agent fera face à des circonstances aléatoires déterminant son coût d'effort. Ces circonstances sont indépendantes d'une relation à l'autre.

La seule différence entre les secteurs est d'être ou non connectés à des échanges d'information. Dans les n premiers secteurs, les sanctions imposées aux agents donnent lieu à un signal transmis au Système II. Dans les $K - n$ autres secteurs, les sanctions sont une affaire privée ne générant aucune information communiquée à des tiers. Les situations analysées dans la section précédente correspondent à $n = 0$ et $n = 1$. Lorsque $n = 0$, aucun secteur n'est connecté et les incitations réputationnelles ne jouent pas de rôle. À l'équilibre, chaque secteur se caractérise alors par la probabilité d'audit p_0 et le surplus S_0 , tels que dérivés précédemment. Lorsque $n = 1$, seul le premier secteur est connecté. À l'équilibre, les relations contractuelles dans ce secteur se caractérisent par p_1 et S_1 tels que dérivés précédemment ; les autres secteurs, non connectés, par p_0 et S_0 .

Étant donné un nombre n quelconque de secteurs connectés, le comportement d'un agent génère un vecteur de signaux (X_1, \dots, X_n) , où X_k est la note obtenue par l'agent dans le secteur k . Les croyances dans le Système II sont donc une fonction $\mu(X_1, \dots, X_n)$. Les incitations réputationnelles dépendent ainsi du nombre de secteurs connectés et des équilibres des relations contractuelles dans ces secteurs. Par symétrie, les valeurs à l'équilibre sont les mêmes dans chaque secteur connecté. Je dénote par p_n la probabilité d'audit à l'équilibre et par S_n le surplus de la relation, où l'indice suggère que ces valeurs dépendent du nombre de secteurs connectés. Pour les $K - n$ secteurs non connectés, rien ne change et les relations contractuelles à l'équilibre sont toujours définies par p_0 et S_0 .

La Proposition 2 a montré que la connexion du secteur 1 est bénéfique aux relations contractuelles dans ce secteur. Elle bénéficie aussi aux relations contractuelles dans le Système II ou, à la limite, n'a que des effets redistributifs. Du point de vue de la société, la connexion du secteur 1 est donc toujours bénéfique. Qu'en est-il lorsqu'on augmente le nombre de secteurs connectés ? On voudra voir comment S_n varie quand n augmente. De plus, les effets sur le Système II seront différents. L'information prédictive obtenue par le Système II pourrait être plus riche qu'avec un seul secteur connecté.

imposer des sanctions inférieures à la sanction maximale admissible, dénotée s ici, avec en revanche un investissement plus important en détection.

Stratégie des agents. Une mauvaise note révèle parfaitement que l'agent est du mauvais type, quel que soit le secteur où elle a été obtenue. L'information communiquée par le vecteur (X_1, \dots, X_n) peut donc se réduire à un signal binaire dénoté $Z_n \in \{M_n, N_n\}$, où M_n signifie que l'agent a été mal noté dans au moins un secteur et N_n qu'il n'a jamais été mal noté. Soient μ_{M_n} et μ_{N_n} les probabilités a posteriori d'avoir affaire à un bon type conditionnellement à M_n et N_n respectivement. Si g_c est le seuil définissant la stratégie des mauvais agents dans chaque secteur connecté et p la probabilité d'audit dans chacun de ces secteurs¹⁵, les probabilités a posteriori satisfont

$$\mu_{M_n} = 0, \quad (13)$$

$$\mu_{N_n} = \frac{\lambda}{\lambda + (1 - \lambda)[1 - p + pF(g_c)]^n}. \quad (14)$$

Le dénominateur de (14) est la probabilité qu'un agent ne soit pas mal noté, sur l'ensemble de la population d'agents et sur l'ensemble des n relations contractuelles connectées. L'équation définit une fonction $\mu_{N_n}(p, g_c)$ croissante en p et décroissante en g_c . En particulier, pour $g_c = h$, on a $\mu_{N_n} = \lambda$ quel que soit p , comme à la Section 2.

J'analyse maintenant la stratégie de meilleure réponse des mauvais agents dans les secteurs connectés, étant donné la valeur de μ_{N_n} . Je suppose que les agents n'ont qu'une capacité limitée de coordination de leurs actions entre les différentes relations contractuelles où ils sont impliqués. L'interprétation est qu'un agent détermine ses actions dans chaque secteur en fonction des circonstances qu'il rencontre dans ce secteur, sans connaître ses circonstances dans les autres secteurs mais en connaissant sa stratégie dans ces autres secteurs.

Considérons l'agent i . Son problème consiste à choisir des seuils g_k^i , où $k = 1, \dots, n$, tels qu'il fera l'effort dans le secteur k si son coût d'effort dans ce secteur satisfait $g \leq g_k^i$ et ne fera pas l'effort si $g > g_k^i$. Le gain attendu de cette stratégie de seuils, sur l'ensemble des secteurs connectés, est

$$\begin{aligned} \bar{u}^i \equiv & - \sum_{k=1}^n \left\{ \int_0^{g_k^i} g f(g) dg + p(1 - F(g_k^i))s \right\} \\ & + \left[\prod_{k=1}^n (1 - p + pF(g_k^i)) \right] v(\mu_{N_n}) \\ & + \left[1 - \prod_{k=1}^n (1 - p + pF(g_k^i)) \right] v(0). \end{aligned} \quad (15)$$

La première ligne est l'espérance de la somme des coûts d'effort et de sanctions sur l'ensemble des secteurs connectés sous la stratégie en question. Les deux

15. On déterminera plus loin les valeurs d'équilibre.

autres lignes donne l'espérance du gain réputationnel, sachant que la probabilité que l'agent ne soit jamais sanctionné est égale à

$$\prod_{k=1}^n (1 - p + pF(g_k^i)).$$

Une solution intérieure à la maximisation de \bar{u}^i , avec $g_k^i < h$ pour tout $k = 1, \dots, n$, satisfait les conditions de premier ordre

$$g_k^i = p \left[s + (v(\mu_{N_n}) - v(0)) \prod_{k' \neq k} (1 - p + pF(g_{k'}^i)) \right], \quad k = 1, \dots, n. \quad (16)$$

La solution symétrique, dénotée g^i , est solution de

$$g^i = p \{ s + [1 - p + pF(g^i)]^{n-1} [v(\mu_{N_n}) - v(0)] \} \quad (17)$$

si cette valeur satisfait $g^i < h$, sinon $g_i = h$. Dans l'équation (17), l'expression

$$[1 - p + pF(g^i)]^{n-1} [v(\mu_{N_n}) - v(0)]$$

représente les incitations réputationnelles de l'agent i dans chaque secteur connecté, étant donné son comportement dans les autres secteurs connectés.

Pour déchiffrer (17), il peut être utile d'analyser le problème de l'agent d'un point de vue *ex post* dans un secteur connecté quelconque, disons dans le secteur 1, étant donné sa stratégie dans les autres secteurs. Pour simplifier l'écriture, posons

$$\pi_i \equiv 1 - p + pF(g^i),$$

ce qui représente la probabilité que l'agent en question ne soit pas mal noté dans un secteur connecté quelconque. Si l'agent fait l'effort dans le secteur 1 lorsque son coût d'effort est g , le gain attendu est

$$-g + \pi_i^{n-1} v(\mu_{N_n}) + (1 - \pi_i^{n-1}) v(0).$$

En effet, l'agent supporte alors le coût d'effort et ne sera pas mal noté dans le secteur 1, mais peut l'être dans d'autres secteurs. La probabilité qu'il ne le soit pas dans aucun des $n - 1$ autres secteurs est π_i^{n-1} , son utilité réputationnelle étant alors $v(\mu_{N_n})$. S'il est mal noté dans au moins un des autres secteurs, son utilité réputationnelle est $v(0)$. Si maintenant l'agent ne fait pas l'effort dans le secteur 1, il sera sanctionné dans ce secteur avec une probabilité p . Son gain attendu est donc

$$p(v(0) - s) + (1 - p)[\pi_i^{n-1} v(\mu_{N_n}) + (1 - \pi_i^{n-1}) v(0)].$$

En comparant les deux gains, l'agent préfère faire l'effort si

$$g \leq p\{s + \pi_i^{n-1}[v(\mu_{N_n}) - v(0)]\},$$

ce qui correspond à la condition (17) ci-dessus.¹⁶

Équilibre. Les principaux adoptent la même stratégie d'audit dans tous les secteurs connectés et les mauvais agents adoptent la même stratégie dans chacune de leurs relations contractuelles dans ces secteurs. En répliquant la démarche de la Section 2, l'équilibre dans les secteurs connectés est une solution en g_c , Δ_n et p des équations :

$$g_c = \min[p(s + \Delta_n), h], \quad (18)$$

$$\Delta_n = [1 - p + pF(g_c)]^{n-1}[v(\mu_{N_n}(p, g_c)) - v(0)], \quad (19)$$

$$c'(p) = (1 - \lambda)\{h - \min[p(s + \Delta_n), h]\}f(\min[p(s + \Delta_n), h])(s + \Delta_n). \quad (20)$$

Ces équations ne diffèrent du système (6), (7) et (8) de la section précédente que par l'équation (19) déterminant les incitations réputationnelles.

Examinons cette équation. Premièrement, pour $n = 1$, (19) se réduit à (7). Dans la notation utilisée ici, les incitations réputationnelles de la section précédente correspondent donc à Δ_1 . Deuxièmement, si $g_c = h$, on a $\Delta_n = v(\lambda) - v(0)$, ce qui définit le niveau des incitations réputationnelles quand tous les agents se comportent efficacement. Le résultat est le même qu'avec un seul secteur connecté. Cela suggère qu'un optimum de premier rang n'est possible que sous les conditions décrites dans la Proposition 1. La multiplication des sources de signaux n'améliore donc pas la possibilité d'un tel équilibre. Troisièmement, Δ_n n'est pas nécessairement décroissant en g_c , contrairement à ce qu'on avait avec un seul secteur connecté.¹⁷ En conséquence, on pourrait alors avoir une multiplicité d'équilibres. Enfin, pour tout $g_c < h$ donné et n grand, Δ_n sera petit. La multiplication des sources de signaux pourrait donc diluer les incitations réputationnelles.

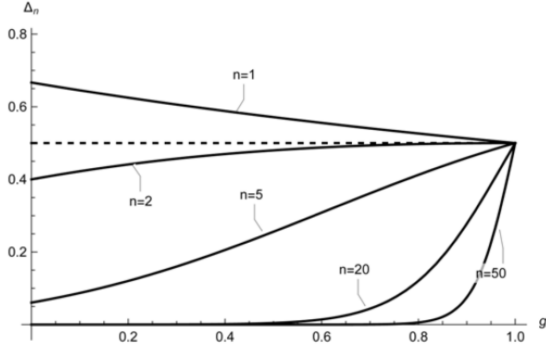
Le graphique 2 présente un exemple numérique des effets du nombre de secteurs connectés sur les incitations réputationnelles, pour toute valeur de g_c . La droite en pointillée indique le niveau de $v(\lambda) - v(0)$, les incitations réputationnelles quand tous les mauvais agents se comportent efficacement. Dans la spéci-

16. La stratégie de l'agent est une stratégie *multiselves* telle que définie dans Piccione et Rubinstein (1997) et Aumann *et al.* (1997). La maximisation de \bar{u}^i donne la stratégie optimale *ex ante*. Par ailleurs, chaque g_k^i est une meilleure réponse aux seuils $g_{k'}^i$, $k' \neq k$, définissant le comportement de l'agent dans les autres ensembles d'information où il pourrait se trouver. Voir Lambert *et al.* (2019) pour une synthèse récente.

17. Pour $n > 1$, le premier facteur dans (19) est croissant en g_c , le second est décroissant.

GRAPHIQUE 2

INCITATIONS RÉPUTATIONNELLES EN FONCTION DE n ET g_c
 $(F(g_c) = g_c, h = 1, p = \lambda = \frac{1}{2}$ ET $v(\mu) = \mu)$



figuration retenue, les courbes Δ_n pour $n > 1$ sont toujours en dessous de la courbe Δ_1 .

Proposition 3. *Lorsqu'il y a un nombre suffisamment grand de secteurs connectés, (i) si $\bar{p}[s + v(\lambda) - v(0)] \geq h$, il y a au moins deux équilibres dont un avec $p_n = \bar{p}$ et $g_c = h$ et un autre avec $p_n > \bar{p}$ et $g_c < h$; (ii) si $\bar{p}[s + v(\lambda) - v(0)] < h$, tout équilibre se caractérise par $p_n > \bar{p}$ et $g_c < h$ et par un surplus de relation $S_n < S_1$.*

Démonstration. Je discute plus loin des équilibres multiples pour le cas (i) et me limite ici à montrer que $S_n < S_1$ sauf lorsque l'optimum de premier rang est atteint. Pour n grand et tout $g_c < h$ et $p > 0$,

$$[1 - p + pF(g_c)]^{n-1} < \frac{v(\lambda) - v(0)}{v(1) - v(0)}$$

puisque l'expression à gauche peut être rendue arbitrairement petite en augmentant n . Comme $\mu_{N_n}(p, g_c) < 1$, il suit que Δ_n défini par (19) satisfait

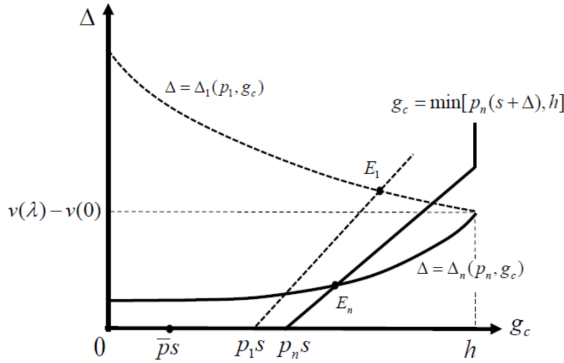
$$\Delta_n < [1 - p + pF(g_c)]^{n-1} [v(1) - v(0)] < v(\lambda) - v(0) \leq \Delta_1$$

La dernière inégalité à droite est satisfaite dans tout équilibre avec un seul secteur connecté. En reprenant l'argument dans la démonstration de la Proposition 1, $\Delta_1 > \Delta_n$ implique $S_1 > S_n$. ■

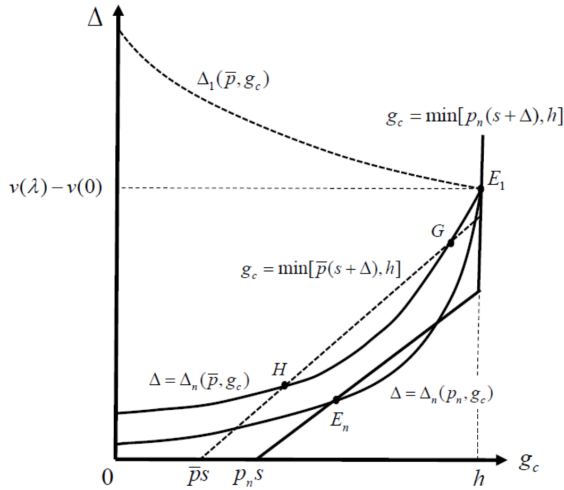
Les Figures 3a et 3b illustrent des équilibres possibles lorsque n est grand. Les incitations réputationnelles $\Delta_n(p, g_c)$ sont définies par (19). Dans le graphique 3a, l'équilibre est E_n . Les principaux ont une stratégie d'audit active et l'optimum de premier rang n'est pas atteint. Les courbes en pointillé reproduisent celles du gra-

phique 1a lorsqu'un seul secteur est connecté. L'équilibre serait alors E_1 . Cet équilibre est évidemment préférable à E_n en ce qui concerne le surplus dans chaque secteur connecté.

GRAPHIQUE 3
ÉQUILIBRES EN E_1 ET E_n



(a) Équilibre en E_n pour $n > 1$ et $\bar{p}(s + v(\lambda) - v(0)) < h$.



(b) Équilibre en E_1 et E_n pour $n > 1$ et $\bar{p}(s + v(\lambda) - v(0)) \geq h$.

Dans le graphique 3b, l'optimum de premier rang en E_1 est toujours un équilibre, quel que soit $n \geq 1$. Il s'agit de l'unique équilibre quand il n'y a qu'un seul

secteur connecté. Pour n grand, cependant, la fonction de seuil pour la probabilité d'audit $p = \bar{p}$ croise la courbe $\Delta_n(\bar{p}, g_c)$ en plusieurs points. Chacune de ces intersections est un équilibre réputationnel pour la stratégie d'audit passive.¹⁸ Toutefois, cette stratégie n'est pas elle-même une stratégie d'équilibre lorsqu'elle induit $g_c < h$. En prenant en compte l'endogénéité de la probabilité d'audit, l'équilibre est donné par le point E_n . La probabilité d'audit est alors $p_n > \bar{p}$, ce qui induit une autre courbe d'incitations réputationnelles.¹⁹ Pour résumer, les deux équilibres possibles sont donc E_1 et E_n . Par conséquent, lorsque l'optimum de premier rang est atteint avec un seul secteur connecté, ce ne serait plus nécessairement le cas avec plusieurs secteurs connectés.

Effets comparés sur les systèmes. Dans le Système I, un secteur intégré aux échanges d'information fait mieux qu'un secteur non connecté. Dans une perspective dynamique, de plus en plus de secteurs voudront donc faire de même. Dans un premier temps, pour n encore petit (et en excluant le cas où l'optimum de premier rang est atteint), l'intégration d'un secteur additionnel pourrait être à l'avantage des secteurs déjà connectés, une externalité positive.²⁰ La Proposition 3 montre, cependant, que les externalités informationnelles négatives sur les secteurs déjà connectés en viendront à dominer. Néanmoins, puisque les incitations réputationnelles augmentent le surplus des relations, il pourrait être avantageux pour le Système I pris globalement que tous les secteurs participent aux échanges d'information. La proposition qui suit montre que ce n'est pas le cas lorsque le nombre potentiel de secteurs accédant aux échanges d'information est très élevé.

Par ailleurs, une évaluation globale devrait aussi incorporer les gains pour les relations contractuelles du Système II. On s'attend à ce que la multiplication des sources de signaux se traduise par une amélioration de l'information transmise à ce système.

Proposition 4. *Pour n très grand et en excluant le cas où l'optimum de premier rang est atteint : (i) dans le Système I, les incitations réputationnelles deviennent négligeables, p_n et S_n tendent vers p_0 et S_0 ; (ii) le signal (X_1, \dots, X_n) tend vers l'information parfaite et en moyenne le surplus des relations dans le Système II tend vers $\lambda Q(1) + (1 - \lambda)Q(0)$.*

Démonstration. Pour le cas où $\bar{p}[s + v(\lambda) - v(0)] < h$, l'argument dans la démonstration de la Proposition 3 montre que Δ_n devient négligeable pour n très grand, ce qui suffit à démontrer (i). Pour démontrer (ii), on utilise le fait que $Z_n \in \{M_n, N_n\}$, tel que défini dans le texte, est une statistique exhaustive de (X_1, \dots, X_n) relativement au type de l'agent. Le surplus des relations dans le Système II est donc

18. Le point G est un équilibre instable, de sorte qu'on pourrait ne retenir que H et E_1 comme équilibres réputationnels possibles, pour cette probabilité d'audit.

19. On peut montrer que, dans la situation représentée au graphique 3b, la courbe $\Delta_n(p_n, g_c)$ est en-dessous de $\Delta_n(\bar{p}, g_c)$.

20. Ce ne serait pas le cas avec la spécification du graphique 2.

$$P_{N_n}Q(\mu_{N_n}) + (1 - P_{N_n})Q(0)$$

où μ_{N_n} satisfait (14) et où

$$P_{N_n} = \lambda + (1 - \lambda)[1 - p + pF(g_c)]^n.$$

Pour tout $g_c < h$ et $p > 0$, P_{N_n} tend vers λ pour n très grand et μ_{N_n} tend vers 1, ce qui signifie que le signal devient parfaitement informatif, d'où l'expression du surplus en (ii). ■

Considérons le surplus moyen des relations dans le Système I sur l'ensemble des secteurs, connectés ou non. Si aucun secteur n'est connecté, le surplus moyen est S_0 . Si quelques secteurs seulement sont intégrés aux échanges d'information, le surplus moyen augmente puisque les secteurs connectés font mieux que ceux qui ne le sont pas et pour lesquels le surplus reste égal à S_0 . En revanche, si le nombre de secteurs connectés devient très grand, le surplus moyen sur l'ensemble des secteurs se rapprochera à nouveau de S_0 . Lorsque le nombre de secteurs connectés a le potentiel de devenir très grand, il y a donc un nombre critique n au-delà duquel l'intégration de secteurs additionnels dans les échanges d'information a un effet globalement négatif sur le Système I.

Si la notation des agents a des effets purement redistributifs dans le Système II, l'évaluation sociale des gains nets des échanges d'information se limite aux considérations précédentes. Si l'information a une valeur productive dans le Système II, les conclusions seraient évidemment modifiées. Il reste qu'un régulateur cherchant à optimiser les échanges d'information devra généralement arbitrer entre les gains en termes incitatifs dans le Système I et les gains de valeur prédictive dans le Système II.

CONCLUSION

Une littérature récente montre qu'une plus grande visibilité des comportements individuels se traduit par des effets incitatifs importants, en accentuant le rôle des motivations réputationnelles, ce qui peut être bénéfique d'un point de vue social. En contrepartie, une plus grande visibilité peut nuire à l'agrégation de l'information relative aux caractéristiques des agents, un effet souvent indésirable.

Un phénomène de ce genre est présent dans l'analyse développée ici. Cependant, le coeur de l'analyse consiste à montrer qu'on peut tout aussi bien obtenir l'effet inverse, selon la façon dont on augmente la visibilité des comportements. J'ai considéré la possibilité qu'un individu doive choisir ses actions simultanément dans un grand nombre de situations, modélisées comme des relations avec des principaux (locateur, employeur, prêteur, etc.). La visibilité du comportement de l'individu repose alors sur le nombre de situations donnant lieu à des évaluations communiquées à des tiers. Ainsi, les nouvelles technologies de type *Big Data*

pourraient permettre le traçage et la consolidation des notes ou scores obtenus par un individu dans plusieurs situations.

Je montre que la multiplicité des sources de notation se traduit, en fait, par une dilution des effets réputationnels. La valeur incitative du score obtenu dans une situation en particulier est plus faible, parce que la réputation dépend aussi du score dans toutes les autres situations impliquant l'individu. En revanche, la qualité de l'information sur les caractéristiques de l'individu est augmentée. L'arbitrage entre incitations et valeur prédictive de l'information opère donc en sens inverse de l'effet mis en évidence dans la littérature récente. Dans la situation que j'ai analysée, une réglementation limitant les échanges de données pourrait se justifier par le souci de préserver les incitations réputationnelles.

BIBLIOGRAPHIE

- ACQUISTI, A., C. TAYLOR et L. WAGMAN (2016) : « The economics of privacy », *Journal of Economic Literature*, 54(2), 442–92.
- AKERLOF, G. A. et R. E. KRANTON (2005) : « Identity and the Economics of Organizations », *Journal of Economic Perspectives*, 19(1), 9–32.
- ALI, S. N. et R. BÉNABOU (2016) : « Image versus information : Changing societal norms and optimal privacy », Document de Travail, National Bureau of Economic Research.
- ARROW, K. (1973) : « The theory of discrimination », dans *Discrimination in Labor Markets*, ed. O. Ashenfelter, et A. Rees. Princeton University Press.
- AUMANN, R. J., S. HART et M. PERRY (1997) : « The absent-minded driver », *Games and Economic Behavior*, 20(1), 102–116.
- BAROCAS, S. et A. D. SELBST (2016) : « Big data's disparate impact », *California Law Review*, 104, 671.
- BERNHEIM, B. D. (1994) : « A theory of conformity », *Journal of Political Economy*, 102(5), 841–877.
- BESLEY, T. et M. GHATAK (2005) : « Competition and incentives with motivated agents », *American Economic Review*, 95(3), 616–636.
- BÉNABOU, R. et J. TIROLE (2003) : « Intrinsic and extrinsic motivation », *Review of Economic Studies*, 70(3), 489–520.
- (2006) : « Incentives and prosocial behavior », *American Economic Review*, 96(5), 1652–1678.
- CHARNESS, G. et M. RABIN (2002) : « Understanding social preferences with simple tests », *The Quarterly Journal of Economics*, 117(3), 817–869.
- CHARNESS, G., L. RIGOTTI et A. RUSTICHINI (2016) : « Social surplus determines cooperation rates in the one-shot Prisoner's Dilemma », *Games and Economic Behavior*, 100, 113–124.
- CHO, I.-K. et D. M. KREPS (1987) : « Signaling games and stable equilibria », *The Quarterly Journal of Economics*, 102(2), 179–221.

- CITRON, D. K. et F. PASQUALE (2014) : « The scored society : Due process for automated predictions », *Washington Law Review*, 89, 1.
- CREEMERS, R. (2018) : « China's social credit system : an evolving practice of control », *Disponible à SSRN 3175792*.
- CROCKER, K. J. et A. SNOW (1986) : « The efficiency effects of categorical discrimination in the insurance industry », *Journal of Political Economy*, 94(2), 321–344.
- D'ALESSANDRO, B., C. O'NEIL et T. LAGATTA (2017) : « Conscientious classification : A data scientist's guide to discrimination-aware classification », *Big Data*, 5(2), 120–134.
- DAUGHETY, A. F. et J. F. REINGANUM (2010) : « Public goods, social pressure, and the choice between privacy and publicity », *American Economic Journal : Microeconomics*, 2(2), 191–221.
- DEFFAINS, B., R. ESPINOSA et C. FLUET (2019) : « Laws and norms : Experimental evidence with liability rules », *International Review of Law and Economics*, 60.
- DEFFAINS, B. et C. FLUET (2020) : « Social norms and legal design », *The Journal of Law, Economics, and Organization*, 36(1), 139–169.
- DEMOUGIN, D. et C. FLUET (2001) : « Monitoring versus incentives », *European Economic Review*, 45(9), 1741–1764.
- FANG, H. et A. MORO (2011) : « Theories of statistical discrimination and affirmative action : A survey », dans *Handbook of social economics*, vol. 1, pp. 133–200. Elsevier.
- FEDERAL TRADE COMMISSION (2016) : « Big Data. A Tool for Inclusion or Exclusion », FTC Report.
- FOX, J. et R. VAN WEELDEN (2012) : « Costly transparency », *Journal of Public Economics*, 96(1-2), 142–150.
- GANUZA, J.-J. et J. S. PENALVA (2010) : « Signal orderings based on dispersion and the supply of private information in auctions », *Econometrica*, 78(3), 1007–1030.
- GOLLIER, C. (2001) : *The Economics of Risk and Time*. The MIT Press.
- HÖLMSTROM, B. (1979) : « Moral hazard and observability », *The Bell Journal of Economics*, pp. 74–91.
- JANN, O. et C. SCHOTTMÜLLER (2020) : « An informational theory of privacy », *The Economic Journal*, 130(625), 93–124.
- LAMBERT, N. S., A. MARPLE et Y. SHOHAM (2019) : « On equilibria in games with imperfect recall », *Games and Economic Behavior*, 113, 164–185.
- MAZYAKI, A. et J. VAN DER WEELE (2019) : « On esteem-based incentives », *International Review of Law and Economics*, 60, 105848.
- MORRIS, S. (2001) : « Political correctness », *Journal of Political Economy*, 109(2), 231–265.

- PHELPS, E. S. ET AL. (1972) : « The Statistical Theory of Racism and Sexism », *American Economic Review*, 62(4), 659–661.
- PICCIONE, M. et A. RUBINSTEIN (1997) : « On the interpretation of decision problems with imperfect recall », *Games and Economic Behavior*, 20(1), 3–24.
- POSNER, R. A. (1977) : « The Right of Privacy », *Georgia Law Review*, 12, 393.
- (1981) : « The Economics of Privacy », *American Economic Review*, 71, 405.
- RAMALINGAM, A. et M. RAUH (2010) : « The Firm as a Socialization Device », *Management Science*, 56(12), 2191–2206.
- ROMEI, A. et S. RUGGIERI (2014) : « A multidisciplinary survey on discrimination analysis », *The Knowledge Engineering Review*, 29(5), 582–638.
- STRAHILEVITZ, L. (2008) : « Reputation Nation : Law in an Era of Ubiquitous Personal Information », *Northwestern University Law Review*, 102(4).
- ŽLIOBAITĖ, I. (2017) : « Measuring discrimination in algorithmic decision making », *Data Mining and Knowledge Discovery*, 31(4), 1060–1089.