L'économie des dispositifs de vérification de l'information : une approche expérimentale

Rémi Le Gall

▶ **To cite this version:**

Rémi Le Gall. L'économie des dispositifs de vérification de l'information : une approche expérimentale. Economies et finances. Université Paris-Est, 2018. Français. NNT : 2018PESC0014 . tel-02152377

HAL Id: tel-02152377
https://theses.hal.science/tel-02152377

Submitted on 11 Jun 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITÉ PARIS-EST

ÉCOLE DOCTORALE ORGANISATIONS, MARCHÉS, INSTITUTIONS (ED 530)

Thèse pour l'obtention du grade de doctorat

Spécialité : Sciences économiques

L'économie des dispositifs de vérification de l'information : une approche expérimentale

Présentée et soutenue publiquement par

LE GALL Rémi

le 12 décembre 2018

Dirigée par

M. L'HORTY Yannick – Professeur des Universités – Université Paris-Est Marne-la-Vallée

Composition du Jury

M. LEGENDRE François	Professeur des Universités Université Paris-Est Créteil	Président du Jury
Mme DE LARQUIER Guillemette	Professeure des Universités Université de Lille	Rapporteuse
M. DIAYE Marc-Arthur	Professeur des Universités Université Paris 1 Panthéon-Sorbonne	Rapporteur
M. LANFRANCHI Joseph	Maître de Conférences Université Paris II Panthéon-Assas	Examinateur
Mme THORON Sylvie	Professeure des Universités Université Paris-Est Créteil	Examinatrice
M. ZAMORA Philippe	Directeur de projet « PIC » Direction de l'Animation de la Recherche et des Études Statistiques (DARES)	Examinateur

L'Université Paris-Est n'entend donner aucune approbation ni improbation aux opinions émises dans cette thèse. Ces opinions doivent être considérées comme propres à leur auteur.

Résumé

L'économie des contrats complets prédit qu'au sein d'une relation d'agence d'une organisation productive, en présence d'aléa moral, un dispositif de vérification de l'information permet de répondre à la fois à un problème de coopération entre les individus et à un problème de coordination des activités. Cependant, au lieu de discipliner des comportements opportunistes, ce dispositif peut engendrer des coûts cachés et réduire la motivation intrinsèque des agents à réaliser une activité qui leur a été attribuée. Sous certaines conditions, il génère une réduction de l'activité, et une perte en termes d'allocation ce qui nuit à l'efficacité.

Dans cette thèse de doctorat, nous avons conduit trois expérimentations contrôlées de terrain avec assignation aléatoire qui visaient à modifier les configurations du dispositif de vérification de l'information afin de résoudre un problème organisationnel propre à trois relations d'agence particulières.

Dans notre premier chapitre, nous avons testé l'effet de la variation de la quantité d'informations détenues par les cotisants sur le dispositif de vérification de la déclaration sociale grâce à des messages ciblés contenant des explications sur le pouvoir de contrôle de l'Agence centrale des organismes de sécurité sociale (Acoss) afin de réduire la fraude sociale.

Dans notre deuxième chapitre, nous avons testé la réduction de l'intensité de la surveillance électronique de la performance des conseillers d'un centre d'appels sous-traitants afin d'améliorer leur qualité de vie au travail.

Enfin, dans notre troisième chapitre, nous avons testé l'effet de la négociation contractuelle du dispositif d'évaluation de la participation des étudiants de licence pendant les travaux dirigés afin d'améliorer leur réussite à l'université.

Mots clés : relation d'agence ; contrat ; dispositif de vérification de l'information ; lutte contre la fraude sociale ; recouvrement des cotisations sociales ; surveillance électronique de la performance ; qualité de vie au travail ; centre d'appels ; évaluation formative ; réussite à l'université ; expérimentation contrôlée de terrain avec assignation aléatoire.

Laboratoire de rattachement

Équipe de de Recherche sur l'Utilisation des Données Individuelles en lien avec la Théorie Économique (ERUDITE – EA 437)

Université Paris-Est Marne-la-Vallée

5 Boulevard Descartes – Cité Descartes – Champs sur Marne

77454 Marne-la-Vallée Cedex 2

Abstract

THE ECONOMICS OF INFORMATION CHECK DEVICES: AN EXPERIMENTAL APPROACH

Economics of complete contracts foresees that within an agency relationship of a productive organization, in the presence of moral hazard, an information check device can address both a problem of cooperation between individuals and a problem of coordination of activities. However, instead of disciplining opportunistic behaviours, this device can generate hidden costs and reduce the intrinsic motivation of agents to perform an activity that has been assigned to them. Under certain conditions, it generates a reduction of the outcome, and a loss in terms of allocation, which is detrimental to efficiency.

In this Ph.D. thesis, we conducted three randomized controlled field trials that aimed at modifying the configurations of the information check device to solve an organizational problem specific to three specific agency relationships.

In our first chapter, we tested the effect of varying the amount of information held by contributors on the social reporting verification device through targeted messages containing explanations of the control power of the Agence centrale des organisations de sécurité sociale (Acoss) in order to reduce social fraud.

In our second chapter, we tested the reduction in the intensity of the electronic monitoring of the performance of advisors of an outsourced call centre in order to improve their quality of life at work.

Finally, in our third chapter, we tested the effect of the contractual negotiation of the device which evaluates the participation of undergraduate students during the tutorials in order to improve their success at the university.

Keywords: agency relationship; contract; information check device; fight against social fraud; recovery of social security contributions; electronic performance monitoring; quality of life at work; call centre; formative evaluation; success at the university; randomized controlled trial.

Laboratory affiliation

Équipe de de Recherche sur l'Utilisation des Données Individuelles en lien avec la Théorie Économique (ERUDITE – EA 437)

Université Paris-Est Marne-la-Vallée

5 Boulevard Descartes – Cité Descartes – Champs sur Marne

77454 Marne-la-Vallée Cedex 2

Remerciements

Cette thèse de doctorat clôture un parcours relativement long, ponctué de plaisirs et de doutes. Elle ouvre de nouvelles perspectives, car ce n'est finalement que le début d'expériences, tant personnelles que professionnelles. Bien qu'elle soit l'aboutissement d'un travail de recherche personnel, beaucoup de personnes ont gravité autour d'elle et il est temps de leur exprimer ma gratitude.

Je voudrais commencer par remercier Yannick L'Horty qui a accepté de m'encadrer dans ce parcours doctoral. Ses conseils m'ont permis de forger mon identité de jeune chercheur et à me remettre en question sur la pertinence de certaines réflexions. Son enthousiasme indéfectible et son stakhanovisme de la réponse à appels à projets de recherche m'auront permis de comprendre les rouages du métier d'enseignant-chercheur ainsi que les considérations entourant la direction d'une institution de recherche telle que la Fédération Travail, emploi et politiques publiques (FR TEPP).

C'est également l'occasion pour moi de remercier Florent Fremigacci, qui suit mon parcours dans une ombre bienveillante depuis ma troisième année de Licence et qui a continué de m'accompagner durant ces quatre dernières années. Sa clairvoyance et son esprit critique sont les deux qualités qui m'ont permis d'avancer sur le chemin de la rigueur scientifique. « Pas de recherche de qualité, sans une question de qualité ! ». Je conserverai précieusement cette devise.

Je remercie particulièrement Guillemette de Larquier ainsi que Marc-Arthur Diaye qui ont accepté de tenir le rôle de rapporteurs. Je n'oublie pas de remercier Joseph Lanfranchi, François Legendre, Sylvie Thoron ainsi que Philippe Zamora qui complètent le jury de thèse. Vos remarques, vos commentaires et vos questions enrichiront considérablement et de plusieurs manières, d'un point de vue théorique ou empirique, cette thèse.

Mes remerciements vont également à la direction du laboratoire ERUDITE du mandat 2013-2018 qui a su me faire confiance et qui a accepté ma proposition de sujet de thèse. Bien qu'éloignée de sa problématique initiale, elle n'aurait pas pu voir le jour sans votre accord.

Je ne pourrais pas rédiger ces remerciements sans exprimer ma reconnaissance envers Nathalie Greenan et Isabelle Gillet. Grâce à elles, j'ai été baigné dans la recherche pendant neuf mois à l'ex-Centre d'Études et de l'Emploi (devenu un programme transversal du Conservatoire National des Arts et Métiers, Centre d'Études de l'Emploi et du Travail). En ayant été directement impliqué dans une expérience de recherche d'envergure, il est sans nul doute possible de penser que sans vous, je ne serai pas là où je me trouve aujourd'hui.

Pendant mon doctorat, j'ai travaillé dans les locaux de la plateforme CELESTE de la Fédération de Recherche TEPP. J'y ai rencontré ce que l'on pourrait nommer le « pilier » de ces murs : Fabrice Gilles. Je voudrais te remercier pour tes conseils de lecture, nos discussions sur l'enseignement supérieur et sur la manière de pratiquer l'économie appliquée. Sans toi, ces bureaux manqueraient d'âme.

J'ai également eu l'opportunité de travailler comme chargé de travaux dirigés à l'Unité de Formation et de Recherche d'Économie-Gestion de l'Université Paris-Est Marne-la-Vallée. J'y ai

rencontré des enseignants-chercheurs incroyables et solidaires pour former les étudiants et les accompagner vers la réussite. Je voudrais les remercier de m'avoir montré l'exemple pour instruire les étudiants et leur donner le goût de l'apprentissage : Thomas Barré, Amina Bécheur, Mélika Ben Salem, Pascale Petit et Guillaume Raoult.

J'ai une pensée spéciale pour Hervé Defalvard, qui m'a accueilli au sein du séminaire de la Chaire Économie Sociale et Solidaire, et qui m'a permis de poursuivre mes réflexions sur l'économie et les biens communs.

Évidemment, je pense à ces personnes que l'on croise régulièrement dans un bureau, dont les conversations rythment une journée de travail, qui deviennent des connaissances, puis des amies, avec qui l'on partage les joies et les incertitudes, des déjeuners, des cafés ou des croissants. Melaine, merci d'avoir été vigilant sur les certitudes que je trainais au début de mon parcours et qui m'a permis de considérer beaucoup d'aspects pragmatiques de la recherche et de l'enseignement, sans toi le doctorat n'aurait pas été aussi riche d'expériences. Souleymane, Jërëjëf pour ta sincérité et ton altruisme à toute épreuve, sans toi ce doctorat n'aurait pas eu la même saveur. Céline, Émilie, Marion et Hajar merci pour votre entraide et votre réalisme, sans vous le doctorat n'aurait pas été aussi féminisé.

Il existe une personne sans qui cet ouvrage n'aurait pas pu voir le jour. Je pense évidemment à toi, mais je n'ai pas de mots plus forts que merci pour te remercier du plus profond de mon être. Sans toi, Marion, ma vie et cette thèse seraient des sujets contrefactuels.

Je voudrais conclure mes remerciements à l'ensemble des enseignants que j'ai côtoyé durant mon parcours scolaire et universitaire. Sans eux, je n'aurais pas eu ni le goût de la curiosité, ni l'envie de produire et de transmettre des connaissances. Il est important de reconnaître leur influence sur ma construction personnelle et d'admettre que les institutions de l'Éducation Nationale et de l'Enseignement Supérieur et de la Recherche ont été des moteurs importants qui m'ont conduit à la rédaction de ce manuscrit. Je souhaite exprimer spécialement ma reconnaissance envers Mesdames H. Very, O. Sieper, Goussot, Pasquier, K. Gandriaux, Chamozzi, E. Tovar, et Messieurs J. Picq, M. Védrine, Gaydou, et P. Batifoulier.

Table des matières

Résumé **5**
Abstract **6**
Remerciements **7**
Table des matières **9**
Introduction générale **13**
I. Le paradigme de l'économie de l'information 14
1. Le positionnement de l'économie des organisations 16
2. La théorie des contrats comme base conceptuelle de la coordination et de la coopération dans les organisations 19
II. La place des dispositifs de vérification de l'information dans la théorie des contrats complets 20
1. La résolution du problème d'aléa moral dans la relation d'agence 20
2. La spécificité des dispositifs de vérification de l'information 23
III. La problématique générale : identifier, modifier, améliorer le rôle des dispositifs de vérification de l'information 31
IV. L'expérimentation contrôlée de terrain avec assignation aléatoire 32
V. Les précautions à envisager dans cette recherche et la présentation des chapitres 39
Chapitre 1. Expliquer les conditions du dispositif de contrôle de la fraude sociale pour améliorer le consentement à la contribution sociale **45**
I. Introduction 46
1. Les éléments de cadrage de la fraude sociale 47
2. Les conséquences sociales et économiques provoquées par la fraude sociale 49
3. La place du dispositif de vérification de l'information dans la lutte contre la fraude sociale 50
4. Concilier la mission de recouvrement avec la mission de vérification de la déclaration sociale 54
II. Revue de la littérature 55
1. Le modèle canonique du consentement au prélèvement obligatoire 55
2. Les enrichissements du modèle 58
3. Les méthodes appliquées pour améliorer le consentement au prélèvement 68
III. Hypothèses de recherche 73
IV. Terrain et méthodologie 75
1. Traitement expérimental 75
2. Procédure de tirage au sort 76
3. Sources d'information 76
4. Limites et retraitement de l'échantillon 79
V. Données 79
1. Participants 79
2. Les indicateurs d'intérêts 86
3. Méthode d'identification des effets 89
VI. Résultats 91
1. Estimation non paramétrique de l'effet causal 91
2. Estimation paramétrique de l'effet causal par double différence 101
VII. Interprétation et discussion des résultats 109

1. Communiquer au sujet du dispositif de vérification de la fraude permet de modifier le comportement des contribuables 109
2. L'effet principal de l'expérimentation : la réduction de la création d'emploi temporaire 111
VIII. Conclusion 114
1. Une contribution à l'analyse de la relation de recouvrement des contributions sociales 114
2. Les limites à la portée des résultats 115
3. Des pistes ouvertes pour approfondir le rôle du dispositif de vérification de la déclaration sociale dans le consentement à la contribution sociale 117
Annexes du chapitre 1 119
4. Contenu des messages envoyés aux groupes traités et au groupe témoin 119
5. Test d'équivalence entre l'échantillon initial et l'échantillon final 125
Chapitre 2. Réduire l'intensité de la surveillance électronique des performances des salariés : une expérimentation dans un centre d'appels sous-traitant 127
I. Introduction 128
1. Les éléments de définition de la surveillance des performances au travail 128
2. Un secteur d'activité recourant traditionnellement à la surveillance des performances au travail : les centres d'appels 132
3. L'approche par la qualité de vie au travail des salariés 135
II. Revue de la littérature 137
1. Les effets théoriques ambivalents de la surveillance électronique de la performance 137
2. Les preuves empiriques des effets ambigus de la surveillance électronique de la performance 141
3. Synthèse 145
III. Développement des hypothèses 145
IV. Terrain et méthodologie 146
1. Contexte de l'expérimentation 146
2. Protocole expérimental 152
3. Limites 158
V. Données 158
1. Les participants à l'expérimentation 158
2. Les mesures de la performance 159
3. Les mesures de la qualité de vie au travail 163
4. Stratégie d'identification 164
VI. Résultats 166
1. Sur les campagnes d'appels entrants 166
2. Les campagnes d'appels sortants 167
3. Synthèse 169
VII. Interprétations 170
1. Une déstabilisation des conseillers téléphoniques 170
2. Un surajustement des superviseurs 172
VIII. Conclusion 175
1. Une révélation détournée des effets néfastes de la surveillance électronique 175
2. Les limites de l'expérimentation peuvent restreindre la portée des résultats 177
3. Une ouverture à un approfondissement du changement d'intensité de la surveillance au travail 178
Annexe du chapitre 2 179
4. Questionnaire de mesure de la qualité de vie au travail spécifique à chaque campagne d'appel 179

Chapitre 3. Négocier les règles d'évaluation pour développer l'accompagnement à la réussite des étudiants 183

I. Introduction 184

1. Les inégalités sociales face à l'échec à l'université et les outils de remédiation 184
2. Les limites de l'évaluation des connaissances comme facteur explicatif de l'échec à l'université 187
3. Les marges de manœuvre possibles pour améliorer le dispositif d'évaluation des connaissances 191
4. Problématique de la recherche 192

II. Revue de la littérature 193

1. La modélisation de la relation enseignant-étudiant 193
2. Des pistes ouvertes par la littérature appliquée pour évaluer la négociation des règles d'évaluation des connaissances 206

III. Hypothèses 215

IV. Terrain de recherche 216

1. Contexte de l'expérimentation 216
2. Protocole expérimental 223

V. Données 228

1. Étude de l'échantillon 228
2. Les indicateurs d'intérêt 230
3. Méthode d'identification des effets 231

VI. Résultats 232

1. La réussite aux examens de contrôle continu et terminal 233
2. La réussite en termes de validation du semestre 237

VII. Discussion et interprétations des résultats 240

1. L'ambivalence du traitement pour les étudiants de première année 240
2. La méfiance des étudiants de deuxième année 242
3. Une interprétation générale 244

VIII. Conclusion 246

1. Des apports à l'économie des organisations 247
2. La portée des résultats est limitée par le protocole expérimental 248
3. Cette expérimentation est une première phase d'exploration pour de futures pistes de recherche 250

Annexes du chapitre 3 252

4. Réussite à l'université en France 252
5. Grille d'évaluation de la participation à l'issue de la négociation 254
6. Test d'équivalence des notes de participation 255

Conclusion générale 257

I. Les principaux résultats de cette recherche 258

II. Les principaux apports à l'issue de cette recherche 260

III. Les principales limites de cette recherche 263

IV. Les pistes ouvertes par cette recherche 265

Bibliographie 267

Introduction générale

Une ancienne tradition de recherche dans les sciences sociales en général et dans les sciences économiques en particulier porte sur les comportements coopératifs entre les individus et les groupes d'individus, ainsi que sur les dispositifs permettant d'influencer ces comportements. En effet, la coopération est un déterminant important de la pérennité d'une organisation sociale (Barnard et Andrews, 1938). L'une des conditions de la coopération à l'intérieur d'une organisation est sa capacité à vérifier que ses membres réalisent les activités qui leur ont été assignées (Mintzberg, 1973).

Dans les configurations sociales et économiques, la manifestation de ce phénomène est courante. Par exemple, circuler sur la route à bord d'une automobile est conditionné d'une part par le respect du Code de la route (limitation de vitesse, respect du sens de circulation…) et d'autre part par la délivrance administrative d'un permis d'utilisation d'un véhicule. Les capacités à enfreindre ces règles et à mettre en danger les autres automobilistes sont une source d'incertitude pour les autorités. Elles doivent donc disposer d'un pouvoir de surveillance pour vérifier que les conducteurs respectent les règles.

Bénéficier de l'indemnisation du chômage est assujettis à l'accomplissement d'actes répétés de recherche d'emploi. L'agence de service public d'accompagnement vers l'emploi est dotée d'un outil de contrôle de ces opérations, en recourant par exemple à une enquête téléphonique auprès des allocataires signalés comme potentiellement irrespectueux envers leur contrat d'accompagnement.

La réalisation de chantiers, à titre privé ou à titre public, entraine un contrôle pour vérifier son état d'avancement et la qualité des matériaux utilisés pour construire ou rénover un bâtiment.

Enfin, avant d'accorder un crédit à la consommation, un banquier doit être certain que son client bénéficie d'une capacité suffisante de remboursement. Il dispose alors d'un pouvoir d'investigation pour identifier les sources de solvabilité.

Le point commun entre ces relations repose sur l'existence d'un contrat, implicite ou explicite, entre deux partenaires d'une relation, conclu pour organiser l'activité et répartir les ressources. Au sein de cette relation contractuelle, il semble crucial pour l'un des contractants de tirer parti d'un dispositif de vérification de l'information pour s'assurer du bon comportement de l'autre. Les organisations sont ainsi en quête d'un outil le plus efficace possible, à la fois pour modifier le comportement des individus et pour acquérir des renseignements supplémentaires afin d'améliorer leurs activités d'allocation de ressources.

Dans un premier temps, nous montrerons comment l'analyse des dispositifs de vérification de l'information se situe au cœur d'un tournant paradigmatique de la science économique. Ce tournant a notamment débouché sur l'économie des contrats, le cadre d'analyse dans lequel s'inscrit notre thèse de doctorat.

À l'issue de cette exposition, nous présenterons notre problématique générale dont l'objectif porte sur l'identification des différentes manières d'obtenir des informations sur l'activité des participants à l'organisation et sur l'amélioration de leurs paramétrages pour résoudre des problèmes organisationnels. Notre thèse de doctorat se situant dans une démarche d'expérimentation de terrain, nous en détaillerons les principes ainsi que les avantages et les inconvénients.

I. Le paradigme de l'économie de l'information

Dans un régime de marchés en concurrence parfaite, l'équilibre général, qui correspond à l'égalisation de l'ensemble des offres à l'ensemble des demandes de biens étant donné un système de prix, est obtenu grâce au comportement de maximisation des intérêts personnels des agents s'échangeant des biens et subissant une contrainte de ressources. Cette situation coïncide à un optimum social au sens de Vilfredo Pareto : il n'est plus possible d'accroître l'utilité d'un individu sans réduire celle d'un autre. C'est le premier théorème de l'économie du bien-être.

D'un point de vue théorique, pour atteindre l'équilibre général, il est essentiel de postuler que l'information est parfaite. Toute l'information pertinente à la prise de décision individuelle transite alors par les prix, qui sont des indicateurs de la rareté des biens.

L'information est un outil de coordination des actions, dans la mesure où le fait d'être informé, savoir quelque chose conditionne la prise de décision rationnelle des individus, ces derniers mettant en balance l'ensemble des gains et des pertes associées à un choix.

L'information est engendrée et véhiculée par des dispositifs techniques, qui sont analysés dans la littérature à travers les concepts de systèmes d'information ou de processeurs d'information (Hurwicz et Reiter, 2006). Si nous spécifions un ensemble d'états de la nature, un ensemble de signaux et un ensemble de décisions, alors il nous est possible d'identifier des systèmes d'information spécifiques qui associent des classes d'états de la nature à des classes d'actions grâce à un vecteur de signaux.

Il est possible de distinguer trois phases dans les processeurs d'information. L'investigation, tout d'abord, produit de l'information à partir d'évènements perçus dans un environnement donné en filtrant les plus pertinents et transforme des données en signaux codifiées pour le transfert d'information. La transmission, ensuite, synthétise les données et doit permettre de les véhiculer

sans distorsions, ce qui requiert une forme de décodage du signal perçu grâce à un choix de langage de communication, de supports physiques et un facteur humain dont l'efficacité ne peut qu'être dépendante de la relation entre les individus. La décision, enfin, intervient après une étape d'interprétation du signal. Cette ultime phase du processeur d'information agit également en rétroaction sur l'ensemble des étapes, parce que la décision est elle-même un élément déclencheur d'évènements.

Dans la théorie de l'équilibre général, le fait de postuler un accès parfait à l'information signifie d'une part que l'information est immédiatement disponible, c'est-à-dire qu'elle est libre d'accès en termes d'acquisition et de compréhension des indications, qu'elle est gratuite et d'autre part, que les individus possèdent la faculté d'obtenir l'intégralité des renseignements disponibles pour prendre leurs décisions.
Cette conception entre en conflit à la fois avec les caractéristiques intrinsèques de l'information et avec le comportement des individus dont les choix et les actions individuelles convoient par ailleurs de l'information. Les individus le savent et cela modifie leur comportement tout en conduisant à des interactions stratégiques.

L'accès à l'information est imparfait parce que les systèmes d'information qui lui servent de support entrainent des distorsions et des coûts. Il y a du « bruit » produit autour du signal, c'est-à-dire une perturbation indésirable superposée au signal utile véhiculant l'information. Par conséquent, l'accès à l'information est partiel.
De plus, l'information est coûteuse, notamment lors des phases de codage, de transport et de décodage de l'information. C'est un coût lié à l'investissement dans des dispositifs techniques et physiques qui gèrent la transmission de l'information.

La possibilité d'atteindre l'équilibre général entre également en conflit avec les caractéristiques cognitives des individus. Dans le processus de détermination de leurs choix et leurs actions, en considérant que les individus sont rationnels, il est possible que leur rationalité soit limitée (Simon, 1959). Leur mémoire peut être défaillante, c'est-à-dire que la capacité à contenir l'ensemble de l'information peut être restreinte. Leur capacité d'identification des variations à l'origine des bruits observés sur l'information, c'est-à-dire leur capacité d'investigation, peut être réduite. Enfin, les individus sont hétérogènes dans leur capacité de traitement de l'information et peuvent rencontrer des problèmes d'interprétation du signal reçu.
Il existe également des situations où certains individus détiennent seuls des informations auxquelles les autres n'ont pas directement accès. C'est le phénomène de l'asymétrie

d'information. Cette caractéristique peut être héritée ou bien provenir du processus de répartition des ressources.

En définitive, il existe des coûts d'information et un problème de fiabilité sur l'information acquise. La réduction des bruits par des dispositifs plus complexes de processeurs d'informations peut se traduire par un système qui peut devenir en lui-même coûteux et producteur de bruits, engendrant ainsi un cercle vicieux. Par conséquent, la possibilité de prendre en compte la dimension d'accès imparfait à l'information a été vécue comme un tournant paradigmatique majeur pour la science économique (Stiglitz, 2002)[1].

La prise en compte des imperfections de l'accès à l'information débouche sur la modélisation du comportement de recherche de l'information. Ce processus de prospection est coûteux en ressources, mais sa détention procure une utilité parce qu'elle rentre dans la prise de décisions. Désormais, il est possible de concevoir qu'il existe une variété de moyens permettant d'avoir un accès à l'information pertinente qui ne reposent plus simplement sur les prix. Dans ce cadre, l'information revêt la qualité de bien rare, utile, et échangeable. C'est donc un bien économique. Tel est le paradigme de l'économie de l'information.

1. Le positionnement de l'économie des organisations

Parmi les nombreux cadres d'analyse de la science économique qui ont été modifiés à la suite de cette prise de conscience (citons l'analyse des marchés avec asymétrie d'information Akerlof, 1970, et la théorie du signal Spence, 1973), nous nous inscrivons à la suite des travaux issus de l'économie des organisations (Ménard, 2012 ; Gibbons et Roberts, 2013). L'organisation est un principe, une force d'intégration, qui assure l'unité d'organismes complexes et différenciés en raison de la division du travail, et qui dépasse tous les cadres spécifiques. Pour Arrow (1970), il est possible d'analyser l'économie, au sens de système de répartition des ressources rares en

[1] Ajoutons à l'idée d'imperfection de l'accès à l'information, l'impossibilité de créer un marché de l'information à cause de ses caractéristiques de bien public (Akerlof, 1970). L'information est un bien non rival lorsque l'utilisation de l'information par un agent ne diminue pas les possibilités d'usage de cette même information par d'autres agents. L'information est un bien non exclusif lorsqu'aucun agent ne peut empêcher un individu d'utiliser un tel bien car dès qu'il est produit, il est à la disposition de tous. Par conséquent, elle génère des externalités, c'est-à-dire toute relation directe entre agents ne donnant pas lieu à un échange à travers un mécanisme de marché. Ces propriétés entrainent trois phénomènes remarquables : la valeur de l'information n'est connue de l'acheteur que lorsqu'il l'obtient (Arrow, 1962) ; la qualité de l'information est également incertaine et ne se révèle qu'à l'usage (Shapiro et Varian, 1998) ; la reproductibilité à coût faible, c'est-à-dire que l'information est coûteuse à produire, mais peu coûteuse à reproduire (Thépaut, 2002). Ces phénomènes permettent difficilement d'établir un marché optimal de l'information, que ce soit au niveau de la demande, puisque les individus ne savent pas fixer un prix correspondant à une quantité en l'absence de qualité et de valeur incertaine, et au niveau de l'offre parce que si le coût marginal de production égalise le prix de marché, l'optimalité parétienne conduirait à prix nul pour l'information en tant que bien public, alors même que les coûts fixes sont positifs. En théorie, le bien ne risque donc pas d'être produit.

fonction des besoins illimités, comme une large organisation. Il est ainsi nécessaire de décomposer les différentes strates qui constituent ce concept.

Les « modes d'organisations » concernent la façon dont les acteurs organisent leurs activités de production et d'échange dans le cadre de règles définies par les institutions[2]. Ces arrangements fournissent ainsi les supports qui gouvernent la manière dont les unités économiques (individus, groupes d'individus) structurent leur coopération ou se concurrencent de manière à assurer les transactions qui permettent de tirer un avantage de la division du travail et de la spécialisation. Dans ce cadre, l'analyse du fonctionnement du marché et de ses imperfections est approfondie. Le marché est un mode d'organisation particulier, dont la finalité est de régler des transactions, et plus précisément de gérer les droits d'usage d'un bien ou d'un service.

Or, il existe bien d'autres modes d'organisation, notamment lorsque les coûts de transaction du marché engendrés par les coûts d'information ne permettent pas de régler efficacement l'allocation des ressources. À partir de l'article séminal de Coase (1937), c'est l'entreprise, l'unité économique de production de biens ou de services, qui serait la plus à même de résoudre ce problème. Cette idée est associée à la reformulation de la place de l'entreprise dans l'économie. En effet, dans la présentation de l'équilibre général, l'organisation productive est circonscrite à la figure de l'autoentrepreneur essentialisée par une fonction de production et une fonction de profit. Pourtant, il est possible de caractériser plus finement l'organisation productive en désignant les modalités structurelles d'organisation de la production dans un système où l'économique a gagné en autonomie par rapport aux autres sphères de l'activité sociale. L'enjeu pour l'entreprise est d'utiliser efficacement les ressources rares grâce à la conjonction de la division du travail et de l'extension des marchés. En d'autres termes, à l'intérieur de la « boîte noire » de l'entreprise existent des interactions sociales qui concourent à rendre efficace l'activité de production.

Il est possible de détailler les caractéristiques des organisations productives qui se distinguent du marché par la mise en place de mécanismes spécifiques de coordination, balisant les actions des individus par des règles qui résultent de négociation. Les organisations productives reposent sur une adhésion volontaire et délibérée des participants. De plus, elles sont flexibles, c'est-à-dire qu'elles ont une capacité à modifier les règles internes qui les gouvernent. Elles font aussi appel à des décisions conjointes entre leurs participants. Dans cette définition, il est possible de

[2] Le concept de mode d'organisation renvoie à l'approche internaliste des organisations, modélisée par la distinction de Davis et al. (1971) entre l'environnement institutionnel, les modes d'organisations et les individus. L'environnement institutionnel se définit par l'ensemble des règles stables, abstraites et impersonnelles, intégré dans des lois, des traditions ou des coutumes et associé à des dispositifs et des systèmes destinés à asseoir et mettre en œuvre des schémas de comportement gouvernant les relations entre individus ou groupes d'individus. Stables, les institutions sont inscrites dans la longue durée, c'est-à-dire qu'elles ne changent que lentement. Abstraites, elles surpassent les règles qui ne concerneraient que des actions particulières. Impersonnelles, elles sont identifiées comme identiques par et pour une majorité d'individus. Normatives, elles délimitent les actions acceptables dans une société à un moment de son histoire.

concevoir que les décisions interviennent dans une structure caractérisée par une différence de positionnement des participants à l'organisation, différence qui s'établit par le concept de hiérarchie (Williamson, 1967) : en dernier ressort de la décision, il y a prédominance des individus qui contrôlent les droits de propriété, directement ou par délégation.
En suivant les termes de Ménard (2012), « l'organisation productive est un mode d'organisation qui résulte du choix conscient des parties de coordonner leurs actions sur une base régulière et pour des objectifs spécifiques, coordination fondée sur un dosage entre commandement et coopération, dont l'efficacité détermine les frontières de cet arrangement. » Nous nous situons dans le prolongement de ce cadre d'analyse.

Que nous considérions le mode d'organisation de l'échange, le marché, ou le mode d'organisation de la production, l'entreprise, il existe des dispositifs de coordination de l'activité, c'est-à-dire d'agencement des tâches diverses selon un plan logique, pour une fin déterminée. L'articulation de ces dispositifs détermine l'efficacité ou les contreperformances des décisions prises. Les dispositifs de coordination permettent aux acteurs autonomes d'effectuer des choix susceptibles de dégager un équilibre entre des offres et des demandes. La coordination, quant à elle, porte sur les moyens mobilisés pour rendre compatibles les actions (Gibbons et Roberts, 2010) : les routines, les incitations, l'autorité, et la coercition.
Nous avons précédemment expliqué que les processeurs d'information étaient des dispositifs imparfaits de la coordination de l'activité d'échange sur le marché, par conséquent il est légitime d'assumer que ces mêmes processeurs sont imparfaits pour coordonner l'activité de production et qu'ils engendrent également des coûts d'information, notamment si nous tenons compte des conséquences des avantages informationnels dont bénéficient certains individus par rapport à d'autres. Il est légitime de penser que dans le cadre d'une organisation productive, où les individus sont en étroite relation physique et continue, les mécanismes de stratégie d'interactions soient plus probables, voire davantage porteurs de conséquences négatives pour l'efficacité productive (Laffont et Martimort, 2002). Il est donc nécessaire d'obtenir des schémas de révélation de l'information contraignant les individus informés à limiter la part de rente informationnelle qu'ils tirent des individus sous-informés.
La figure du contrat s'est alors imposée comme l'élément déterminant de l'organisation pour assurer une fonction de coordination complémentaire pour réduire les coûts d'information. Il fait l'objet d'une attention importante de la science économique au même titre que la discipline juridique ou la sociologie pour en définir ses particularités et ses objectifs (Brousseau et Glachant, 2000). L'analyse de la genèse des contrats éclaire la façon dont les agents conçoivent les règles et les structures de prise de décision qui vont encadrer leurs comportements. L'observation de

l'évolution des dispositifs contractuels aide à comprendre comment se transforment les structures d'encadrement de l'activité économique.

2. La théorie des contrats comme base conceptuelle de la coordination et de la coopération dans les organisations

Un contrat est une règle qui part d'un commun accord, explicite ou tacite, par laquelle un ou plusieurs individus s'engagent envers un ou plusieurs autres sur leurs comportements réciproques à donner quelque chose, à faire ou à ne pas faire une activité. Ainsi, une organisation productive peut s'apprécier comme le lieu de contractualisation par excellence, puisque le contrat permet de mettre en commun et d'organiser les activités de plusieurs parties dans le but de produire des biens ou des services.

Il est possible d'attacher au contrat plusieurs critères pour préciser son identification dans les relations organisationnelles. Le critère de complétude permet d'envisager si l'accord, d'une part, comporte des éléments vérifiables par une tierce partie et d'autre part, conduit à des arbitrages en cas de non-respect (Fama et Jensen, 1983) ou s'il repose sur un consentement mutuel des deux parties contractantes (Azariadis, 1975). Ce critère de vérifiabilité permet d'identifier des types de contrats appariés avec des modes d'organisation où les institutions de gouvernance sont clairement identifiées, tandis que son inexistence engendre un continuum d'arrangements bilatéraux ou multilatéraux, sans qu'il soit possible d'identifier parfaitement la place des institutions tierces (Alchian et Demsetz, 1972).

Le critère d'étendue fait référence à la capacité du contrat à tenir plus ou moins compte de la capacité à observer des états du monde tout en leur faisant correspondre des probabilités d'occurrence, ainsi qu'au degré de symétrie de l'accès à l'information entre les deux parties. Ce critère d'observabilité permet de rendre compte des mécanismes de coercition et d'incitation disposés à faire coopérer les parties du contrat. Nous y reviendrons en détail par la suite. En définitive, les contrats contribuent à régulariser l'action des parties dans un univers soumis aux imperfections d'accès à l'information et à l'incertitude.

Parmi les quatre courants de recherche en sciences économiques qui analysent les contrats sous le prisme de ces deux critères, nous nous positionnons dans le cadre de la théorie des contrats complets (Bolton et al., 2005)[3]. Dans ce cadre d'analyse, les individus détiennent une rationalité

[3] Outre celui des contrats complets, il existe trois autres courants dans la théorie des contrats (Brousseau et Glachand, 2000 ; Ménard, 2012). La théorie des contrats contingents est l'application contractualiste à l'approche walrasienne des échanges et de la production. La rationalité des individus est substantielle, tous les états du monde sont associés à des lois de probabilité d'occurrence, l'information détenue par les individus est symétrique et les institutions de vérifications des termes du contrat sont parfaites, dans le sens où elles garantissent l'absence de déviation des contractants. Dans ce cas, l'organisation établit de manière centralisée et simultanée toutes les valeurs et l'allocation des biens et services. Comme pour la théorie des contrats complets, l'analyse porte sur les conditions de

substantive, c'est-à-dire qu'ils disposent d'une information complète sur la structure des problèmes auxquels ils sont confrontés (capacité de calcul infini et ensemble complet de préférences), mais l'information n'est pas distribuée de manière symétrique entre les individus. L'avenir est incertain et aléatoire, ce qui fait intervenir la notion de risque (Knight, 1921). Chaque évènement incertain est associé à une loi de probabilité. Les agents imaginent alors des solutions les plus efficaces en fonction des différents états possibles de la nature et calculent des espérances de résultat. En revanche, tout ce qui est observable est vérifiable, ce qui garantit l'exécution des engagements. Cette hypothèse permet de formuler des schémas de coordination qui révèlent l'information et qui permettent d'inciter à la coopération. Par conséquent, c'est ce cadre qui a permis de fournir un support théorique à l'existence de dispositif de vérification de l'information.

II. La place des dispositifs de vérification de l'information dans la théorie des contrats complets

1. La résolution du problème d'aléa moral dans la relation d'agence

Dans l'organisation productive, dès lors qu'il existe un contrat, un lien de subordination, une asymétrie informationnelle et un risque entre les contractants, il est possible d'invoquer la notion de la relation d'agence. Ce modèle permet de concevoir la capacité pour un contractant de surveiller l'autre contractant afin de parvenir à la coopération au sein de l'organisation.

Le modèle de base considère l'interaction entre deux individus où l'un d'entre eux, l'agent, agit pour, à la place, ou comme un représentant de l'autre désigné comme le principal dans un domaine de décision particulier (Ross, 1973). Il est possible de dépasser la notion de délégation d'activité qui exprime l'idée de « faire faire » à l'agent ce que le principal peut réaliser, pour utiliser l'idée de la répartition des tâches dans l'organisation (Levinthal, 1988). De cette manière, les parties de la relation ont des objectifs et des intérêts différents au sein de l'organisation (March et Simon, 1993). De plus, ces deux individus opèrent dans un environnement incertain, ce qui procure des degrés d'aversion au risque différents entre les deux parties de la relation. Le partage du risque est donc désirable. En général, l'action optimale pour l'agent dépendra de la répartition du risque entre les deux individus (Grossman et Hart, 1983). Dans ce contexte, ce qui rend la

définition *ex ante* (avant la signature) du contrat optimal, ce qui facilite les procédures de modélisation, y compris dans les situations d'incertitude grâce à des clauses de conditionnalité (Tirole, 1999). La limite de cette approche est qu'il est en pratique invraisemblable de pouvoir envisager toutes les situations possibles, et qu'en théorie il existerait des coûts de transactions (Spiller, 2010).

C'est pourquoi deux autres types de contrats ont tenu compte de ces limites : la théorie des contrats incomplets (Grossman et Hart, 1986) où tout peut être observé, mais ne peut pas être systématiquement vérifié, postule l'existence d'institutions imparfaites de vérification de l'information ; la théorie néo-institutionnelle des coûts de transaction repose sur l'hypothèse de rationalité limitée (Simon, 1947) et tient compte des modalités de gestion de l'incomplétude en stipulant des clauses de renégociation, de recours à une tierce partie et à la hiérarchie. Ces deux courants théoriques alternatifs définissent ainsi des dispositifs d'ajustement *ex post* au contrat.

relation « complexe » est l'incertitude du principal quant aux actions et aux caractéristiques de l'agent (Eisenhardt, 1989). Les exemples de relation d'agence sont universels et tous les arrangements contractuels contiennent des éléments d'agence (Ross, 1973).
Si la variable faisant l'objet d'une diffusion asymétrique de l'information est exogène, c'est-à-dire non manipulable par la partie qui la détient, par exemple sur la qualité de l'individu qui tient le rôle de l'agent, les modèles analyseront la situation de sélection adverse. Si la variable est endogène et manipulable par l'agent, alors les modèles analyseront la situation d'aléa moral. Nous nous inscrivons à la suite de ce courant d'analyse.

L'aléa moral survient lorsque les agents ont été sélectionnés pour leurs connaissances spécifiques, mais le principal ne peut jamais complètement vérifier sans coût l'action de l'agent (Arrow, 1968). Or, la réalisation de cette activité est coûteuse pour l'agent (Mirrlees, 1999). L'agent peut ainsi entreprendre des actions qui accroissent son gain personnel, ce qui se traduit théoriquement par une réduction de l'effort, mais cela réduit le surplus total de la relation. Ainsi, nous définissons l'aléa moral comme la variation incertaine d'un évènement lié à une prise de risque de la part de l'agent. Cette situation se situe donc après la contractualisation (Bolton et Dewatripont, 2001).
Les déterminants de ces actions d'aléa moral sont les variables d'effort qui influencent positivement le niveau de production de l'agent, mais créent également une désutilité pour l'agent. Les actions de l'agent peuvent aussi être un ensemble plus complexe de décisions qui définissent les tâches ou les attributions d'activités de l'agent. Celui-ci peut choisir plusieurs projets à les accomplir à la place du principal (Laffont et Martimort, 2002).
In fine, la relation d'agence porteuse d'un aléa moral pose le problème de la coopération au sein de l'organisation. La coopération concerne les comportements et vise à amener les parties à utiliser les moyens de coordination dans les actions qu'ils entreprennent en cohérence avec les objectifs poursuivis. Une coopération est un système par lequel les personnes associées à un but commun ont des droits égaux à la gestion et se répartissent le profit au prorata de leur activité. L'existence d'une asymétrie d'information sur le comportement d'une des deux parties de la relation d'agence conduit à une situation de rupture de la coopération, ce qui peut donc se traduire par une inefficacité organisationnelle.
L'aléa moral ne constituerait pas un problème si le principal et l'agent partageaient les mêmes objectifs. Les coûts d'agences augmentent sous l'aléa moral à cause du conflit sur les actions à entreprendre. Le fait de ne pas pouvoir observer les actions des agents peut empêcher une résolution efficace du conflit d'intérêts puisqu'un contrat ne peut pas stipuler quelles actions entreprendre.

Dans ce cadre, il s'agit de définir quel serait le contrat optimal d'agence qui réduirait l'aléa moral. Le contrat conduirait alors à révéler l'information par des incitations adéquates et/ou à restreindre le domaine de choix des individus. De cette manière, on confère un rôle supplémentaire au contrat, précédemment défini comme outil de coordination des activités de l'organisation, celui de faire coopérer des individus aux intérêts divergents dans une situation d'incertitude.

La théorie des contrats complets permet de déterminer quelle est la compensation minimale que le principal doit offrir à l'agent pour s'assurer que ce dernier accepte de réaliser la tâche parce que ne pas y participer engendre des pertes (contrainte de participation) et qu'il la réalise à un niveau satisfaisant (contrainte d'incitation) qui motive l'agent à adopter un comportement conforme à l'intérêt du principal, lorsqu'il ne peut pas observer ces actions (van Ackere, 1993). Il faut choisir le meilleur contrat étant donné l'incertitude pesant sur le résultat, sur l'aversion au risque de l'agent et du principal et sur la quantité d'informations disponible.

En général, la présence d'asymétrie informationnelle en faveur de l'agent permet de formuler l'hypothèse selon laquelle l'incertitude qui conditionne le montant du résultat issu de l'activité de l'agent dépend de l'effort de ce dernier. En d'autres termes, chaque probabilité des différents états du monde et le volume espéré de production dépendent explicitement de l'effort déployé par l'agent. Une fois que l'incertitude est rendue endogène, le résultat est observé et la compensation, issue du partage du résultat et de la prise de risque de l'agent, est versée à l'agent. Cette hypothèse paraît réaliste à partir du moment où le principal ne peut concevoir un contrat que sur la base des performances observables.

En définitive, le schéma de compensation repose sur une rémunération conditionnée à des signaux résultant du comportement de l'agent (Brousseau et Glachant, 2000). Parmi les contrats possibles, le principal préfère celui qui permet d'induire un effort positif au minimum de coûts.

En reprenant la description initiale de la relation d'agence, le résultat de l'activité de l'agent doit, en principe, être partagé entre les deux parties en fonction de la répartition du risque. Si l'agent est neutre face au risque et a suffisamment de ressources financières, alors le problème de contractualisation a une solution directe, l'agent paie un montant au principal et en retour a tout le profit, c'est-à-dire que l'agent doit prendre tous les risques. Si l'agent a une aversion au risque, il n'est pas optimal pour lui de prendre en charge le risque. Le gain des incitations doit être comparé aux pertes d'un partage du risque sous optimal.

En présence d'aléa moral, l'allocation du risque Pareto-optimale n'est pas atteinte. À la place, le principal doit trouver une allocation du risque qui est appropriée pour l'agent afin qu'il réalise une action. En particulier, si l'agent craint le risque, la solution de premier rang n'est pas atteignable. Dans ce cas, la solution de second rang est obtenue en arbitrant le partage des risques avec une

provision d'incitation. Le paiement de l'agent s'effectue en fonction du résultat, mais il ne laisse pas tout l'agent supporter le risque (Jost, 1991). La réduction du risque imposée à l'agent permet de réduire la compensation et de satisfaire les contraintes de participation. C'est une manière de réduire l'écart entre le « *first best contract* » qui correspond à celui qui est atteignable en situation d'incertitude sans asymétrie d'information et le « *second-best contract* » qui est atteignable en situation d'incertitude et avec asymétrie d'information (Levinthal, 1988).

Cette conception du contrat permettant de réduire l'aléa moral correspond plutôt à ce qui s'observe dans certaines relations contractuelles, notamment les contrats commerciaux ou les contrats de gouvernance. L'effort de l'agent est induit par une rémunération à la performance personnelle. L'agent est incité à réaliser le niveau d'effort souhaité par le principal.
L'incitation est ainsi conçue comme un dispositif fondé sur l'espérance de gain, pour amener les agents à révéler l'information qu'ils détiennent et à agir en conformité aux objectifs fixés sur la base de cette information. Cette conception de l'incitation est porteuse de limites et notamment celle de contrecarrer la motivation intrinsèque à réaliser une tâche (Bénabou et Tirole 2003 et 2006 ; Fehr et Falk, 2002 ; Festré et Garrouste, 2015 ; Frey et Jengen, 2001 ; Gnezzy et al. 2011). La promesse d'une récompense peut être interprétée comme un manque de confiance du principal envers l'agent, parce que ce dernier croit que les incitations explicites sont nécessaires pour induire un effort qu'il aurait réalisé de toute façon, ou pour indiquer la pénibilité de la tâche.
Le rôle dual des incitations a été abondamment analysé d'un point de vue théorique et empirique. Cependant, il existe relativement moins d'apports concernant le support des incitations. Nous avons en effet décrit que le schéma de compensation intervenait à partir du moment où le principal observait le résultat, qui était déterminé par l'effort de l'agent. Le schéma d'incitation repose sur une information vérifiable, observable par un tiers. Or, le fait d'observer le résultat correspond à un dispositif de vérification de l'information. Par conséquent, le contrat contient en lui-même un dispositif de surveillance. Cette dimension est l'objet même de cette thèse de doctorat.

2. La spécificité des dispositifs de vérification de l'information

Alchian et Demsetz (1972) considéraient que l'organisation de la firme est différente du marché, parce qu'il faut « surveiller les activités des individus à l'intérieur de l'organisation ». Leur analyse a permis de développer l'idée selon laquelle les incitations et la capacité de surveillance de l'organisation sont complémentaires. Pratt et Zeckhauser (1985) ont également insisté sur le fait que les pertes de la relation d'agence sont les plus sévères lorsque les intérêts des parties sont divergents et lorsque le contrôle de l'agent est coûteux. En effet, baser les récompenses sur une

observation restreinte du résultat impose un risque sur l'agent en rendant sa compensation corrélée à plusieurs facteurs indépendants de sa volonté et de son contrôle. Il est donc possible d'affiner le modèle de base, en formulant l'hypothèse que le résultat issu de l'activité de l'agent est observable et vérifiable, mais qu'il s'agit d'un indicateur imparfaitement corrélé avec les actions cachées de l'agent. Une mesure idéale de l'effort devrait refléter les seuls facteurs dont l'agent a le contrôle.

Dans la plupart des situations contractuelles, le principal peut acquérir des informations additionnelles sur l'action de l'agent au-delà de son résultat. Cependant, obtenir des connaissances sur ces actions est coûteux et n'est pas toujours possible : même si l'action est observée, il est difficile de la décrire en avance et même si elle peut être décrite et observée, il peut être impossible de vérifier quelles actions ont véritablement été entreprises. La surveillance se réfère donc à la production et à la collecte des informations sur les actions de l'agent.

Dans ce cadre, une structure de surveillance plus détaillée élargit les bases possibles des schémas de compensation. La surveillance permet au principal de distinguer si les variations du résultat sont liées aux actions de l'agent ou liées aux états de la nature. De cette manière, il est possible de concevoir un contrat où le résultat de l'activité n'est plus seulement dépendant de l'effort de l'agent.

Plus généralement, il existe un arbitrage entre imposer un risque à l'agent, le récompenser et prendre en charge les activités de surveillance coûteuse (Demougin et Fluet, 2001). Dans certains cas, l'information est parfaite, lorsque l'agent est assigné à une tâche simple et dont la réalisation est vérifiée aisément, et dans d'autres l'information est imparfaite. Par conséquent, le coût de l'incitation dépend de la qualité de l'information. Puisque l'information est coûteuse, le principal doit considérer le meilleur moyen pour inciter l'agent : si la surveillance est difficile, le principal donne des incitations élevées ; si les incitations ne sont pas possibles et si l'agent est très peu enclin au risque, le principal va davantage surveiller.

Dans la littérature théorique, il existe plusieurs manières d'appréhender le système de surveillance. En général, les méthodes de surveillance dépendent des préférences du principal et de l'agent, de la distribution du lien entre le résultat de l'agent, la surveillance supplémentaire du principal, l'ensemble des actions disponibles pour l'agent et du coût d'acquisition d'une information supplémentaire (Dye, 1986).

Dans un premier temps, il est possible de concevoir un système de surveillance qui repose sur le signal du résultat issu de l'activité de l'agent qui correspond à la surveillance contingente (Hölmstrom, 1979 ; Jost, 1991 ; Demougin et Fluet, 2001 ; Kim, 1995 ; Fagart et Sinclair Desgagné, 2007).

Dans ce modèle, avant de signer le contrat, le principal choisit un système de surveillance et paie le coût associé. Le système de surveillance est alors une connaissance commune de la relation d'agence. Un agent peut faire des erreurs dans son travail, ce qui correspond implicitement à son niveau d'effort : plus il fait des efforts, moins il commet d'erreurs et réciproquement. Le nombre d'erreurs détectées dépend de l'intensité de la surveillance. Une surveillance plus intense génère un signal plus efficient ce qui implique que la précision est croissante avec le coût de la surveillance. Pour inciter à une action au coût le plus bas, le principal doit minimiser l'occurrence des erreurs étant donné le niveau de surveillance et l'effort de l'agent.

Puis, le principal offre un contrat à prendre ou à laisser. L'agent réalise une action qui n'est pas observée. Un signal résulte de l'action qui est observable par le système de surveillance. En définitive, c'est le signal qui détermine la compensation de l'agent, c'est-à-dire qu'elle doit être basée sur les informations mesurant la valeur réalisée du résultat de l'agent.

Cette intuition est le principe de la statistique suffisante ou le théorème d'informativité (Hölmstrom, 1979). Supposons que le principal considère réaliser une compensation liée à un signal. Un principal va préférer une structure d'information à une autre si et seulement si au moins une information supplémentaire donne une précision sur l'action de l'agent. On dit alors que le signal est informatif au sens de Blackwell (1953) sur l'effort de l'agent, c'est-à-dire que le résultat est une statistique suffisante de l'action (Grossman et Hart, 1983 ; Gjesdal, 1982). Un signal qui est corrélé à un bruit a potentiellement plus de valeur pour le principal tandis qu'un signal qui n'est pas corrélé au bruit et à l'action est toujours inutile. Plus généralement, le choix d'un principal entre deux systèmes d'information résulte d'un arbitrage entre les coûts directs de ce système qui accroissent l'utilisation du signal des actions et les possibles bénéfices procurés par ces systèmes pour réduire les coûts d'agence.

Ce principe d'informativité permet de concevoir que l'introduction d'un système de surveillance peut motiver l'agent à faire des actions qui engendrent plus de bien-être pour les deux parties de la relation. Dans ce cas, la surveillance permet d'améliorer la définition du contrat optimal (Jost, 1991).

Dans un second temps, il est possible de concevoir un système de surveillance comme une fonction qui relie les états de la nature et les niveaux d'effort à un signal supplémentaire, c'est-à-dire que l'observation donne de l'information sur l'effort qui n'est pas reflété dans le résultat lui-même (Baiman et Demski, 1980 ; Dye, 1986). Par exemple, c'est prendre la décision de produire des informations additionnelles conditionnellement à des données de performances préliminaires de l'agent. Après avoir observé le résultat, le principal peut obtenir des informations supplémentaires sur l'action de l'agent.

Dans ce cadre, le principal détient un dispositif coûteux, qui lui permet d'observer une variable supplémentaire qui est statistiquement reliée à l'action de l'agent, mais indépendante du résultat après l'avoir observé. Si le principal obtient une information indépendante du résultat, l'agent reçoit une compensation qui peut dépendre des valeurs du résultat et de cette variable.
Le principal est ainsi supposé être capable de réaliser un contrat et une stratégie de surveillance qui spécifient la probabilité qu'il aille examiner les actions des agents comme une fonction de la valeur réalisée du résultat. L'agent réalise son action et une variable aléatoire influence le résultat. Si l'investigation intervient, une variable aléatoire secondaire est obtenue. Par conséquent, l'action de l'agent dépend de la compensation et du dispositif d'investigation offert par le principal.
Dans ce modèle, il n'y a pas de surprises après avoir surveillé les actions de l'agent. Puisque le principal connait les informations sur le problème de décision de l'agent, il connait quelles actions l'agent entreprend, et donc quelles actions il va observer quand l'investigation intervient. Il faudrait donc inclure des coûts de spécifications des contrats et les coûts de communication.
Lorsque le système de surveillance est imparfait et produit de l'incertitude sur le schéma de compensation pour un agent qui a de l'aversion pour le risque, la menace d'être évalué rend indésirable le résultat déclenchant la surveillance, ce qui provoque une incitation pour l'agent de réduire la probabilité d'obtenir ces résultats (Baiman et Demski, 1980). Par conséquent, il est seulement optimal de surveiller l'effort des agents quand le résultat est inférieur à un certain seuil.
Lorsque le principal peut utiliser une surveillance qui révèle toute l'information, le dispositif optimal est celui qui utilise un seuil de résultat bas pour déclencher la surveillance (Dye, 1986). La désutilité associée à l'incertitude de la surveillance est moindre pour un agent qui a moins d'aversion pour le risque. Pour garder la menace de la surveillance effective, des alternatives sont possibles, soit il faut concevoir les pénalités en rendant impossible qu'elles puissent être favorables, soit il faut accroître la probabilité d'être contrôlé. Pour les agents moins enclins au risque, il est optimal de rendre la compensation contingente à l'investigation favorable, ce qui procure à l'agent une envie d'être contrôlé. Pour les agents plus enclins au risque, il est optimal de mettre des sanctions.

Ces deux visions de la surveillance ne sont pas opposées, mais plutôt complémentaires. Pour Demougin et Fluet (2000), leur mise en place dépend du type d'activité demandé à l'agent. En plus du critère d'informativité du système d'observation et de vérification, la littérature a également fourni des indicateurs d'ordonnancement des systèmes d'information (Kim, 1995 ; Fagart et Sinclair-Desgagné, 2005). Kim (1995) a développé, le « *mean preserving spread* » (MPS) qui stipule qu'un système d'information est meilleur quand le principal peut inférer avec une plus grande précision les actions de l'agent. Pour Fagart et Sinclair-Desgagné (2005), un système d'information généré par une surveillance sur une variable indépendante du résultat doit être

classé en fonction d'une « dominance du troisième ordre ». Cela implique des choix de surveillance selon le risque local ajouté, un comportement qui est capturé par le concept de « prudence » de l'agent (Eekoudt et al., 2005), qui est l'attitude consistant à réfléchir à la portée et aux conséquences de ses actes, à prendre ses dispositions pour éviter tout danger, toute erreur, tout risque inutile. Notons que, malgré la rigueur mathématique de ces travaux, il ne semble pas qu'il y ait une volonté de description plus applicable de tels dispositifs de collecte de l'information ni de comparaison avec les pratiques existantes dans les organisations.

Dans la littérature sur le système d'information dans le contrat, nous pouvons également faire référence à la notion de principal informé développée par Tirole (1992) et poursuivie dans le cadre des dispositifs de vérification de l'information par Chade et Silvers (2002) et Silvers (2012). Dans ce modèle, le principal détient une information privée sur le système de vérification du résultat de l'agent et le contrat qu'il propose signale cette situation à l'agent. Ces modèles offrent la possibilité de combiner la sélection adverse sur le type de principal avec l'aléa moral sur l'effort de l'agent. Silvers (2012) révèle, à partir d'une situation *ex ante* (avant que le signal sur le type de principal soit observé), que l'agent préfère que le principal soit en capacité de surveiller son comportement et que cette information soit partagée, tandis que le principal est plus enclin à ne pas détenir d'informations privées en raison de leur coût. En revanche, le principal préfère que le signal sur le dispositif d'information soit connu et partagé à une situation où il n'existe pas d'information. Pour tout équilibre séparant avec information privée, il existe un contrat Pareto-optimal où les deux parties ont une préférence pour l'information partagée sur le dispositif de vérification du résultat.

Pour résumer le propos, quel que soit le signal qui est informatif sur l'effort de l'agent, il doit être inclus comme argument dans son schéma de compensation. Par conséquent, c'est un terme du contrat d'agence. Ces travaux concluent que la surveillance permet d'inciter l'agent à produire l'effort attendu par le principal, ce qui formule l'intuition d'un effet disciplinant de la surveillance sur le comportement de l'agent, y compris si la surveillance est imparfaite et coûteuse pour le principal. D'un point de vue empirique, les effets disciplinant de la surveillance existent et procurent des effets positifs sur l'effort des individus (Nagin et al., 2002 ; Boly, 2011 ; Belot et Schröder, 2015).

Néanmoins, à l'instar des effets ambigus des incitations sur les comportements, que ce soit un effet disciplinant ou un effet d'éviction de la motivation intrinsèque (Frey, 1993), la surveillance est elle-même ambivalente quant aux effets sur le comportement des agents. Chronologiquement, trois arguments ont émergé dans la littérature pour montrer les effets négatifs de la surveillance sur l'effort des agents : l'éviction de la motivation intrinsèque (Frey, 1993) ; la recherche d'une

stratégie de contournement de la règle de surveillance (Cowen et Glazer, 1996) ; les coûts cachés de la surveillance (Falk et Kosfeld, 2004).

Quel que soit le mode d'organisation, les agents adaptent leur comportement aussi bien en fonction des variations de gains que par comparaison avec leur entourage immédiat. En considérant cette idée, on peut mettre en avant la notion de motivation qui repose sur une appréhension du bien-être fondé sur la perception du statut eu égard à la qualification ou le passé éducatif (Pfeffer, 1995). Dans une organisation, les agents se construisent une identité marquée par les réseaux dans lesquels ils opèrent. Il en résulte des motivations conduisant à l'échange de bons procédés ou à des stratégies de résistance passive.

Pour Frey (1993), un contrat psychologique implicite existe entre le principal et l'agent. Ce contrat dépend des bonnes relations avec le principal, que l'on peut apprécier par « l'atmosphère » (Williamson, 1975), la culture de l'organisation (Akerlof et Yellen, 1986), la confiance (Baker, Jensen et Murply, 1988) et le fait de faire des choses dans l'organisation sans récompense (Simon, 1991).

Lorsque le principal fait le choix d'accroître la supervision, ce dernier indique sa méfiance envers la possibilité pour l'agent de réaliser sa tâche. Celui-ci en percevant le signe de méfiance induit par la surveillance peut réduire sa motivation intrinsèque à réaliser l'activité demandée et *in fine* son effort, ce qui se traduit par une rupture du contrat entre les deux parties.

Frey (1993), en s'appuyant sur des considérations psychologiques (Deci et Ryan, 1985), considère qu'il peut exister deux conditions qui conduisent à évincer la motivation : quand « l'auto-détermination » est réduite par le principal, ce qui substitue du contrôle intrinsèque par du contrôle extrinsèque sur la tâche à réaliser ; quand « l'auto-évaluation » est réduite par le principal. Par conséquent, la surveillance, quand elle indique que le principal est convaincu que l'agent ne sait ou ne veut pas remplir la tâche assignée pour sa satisfaction, peut réduire la motivation intrinsèque. Pour que l'effet d'éviction apparaisse, il faut que la relation entre le principal et l'agent soit personnelle, tandis que l'effet disciplinant domine lorsque la relation est impersonnelle. En général, les deux effets coexistent dans la relation d'agence.

Pour modéliser cet effet d'éviction dans une relation d'agence, Frey (1993) considère que le résultat marginal associé à l'effort de l'agent décroit au fur et à mesure que l'intensité de la surveillance augmente, puisqu'il existe une opportunité extérieure à tricher, cet état de fait étant causé par la rupture de la confiance dans la relation avec le principal. D'une certaine manière, la méfiance du principal induit l'agent à être conforme à cette attente de tricherie.

L'effort optimal de l'agent est atteint lorsque le gain marginal de l'effort est égal au coût marginal de l'effort. En différenciant l'effort optimal par rapport à l'intensité de la surveillance, trois cas peuvent apparaître : un effet disciplinant sans effet d'éviction qui augmente l'effort ; un effet

d'éviction sans effet disciplinant qui réduit l'effort ; les deux effets sont présents, mais l'impact sur l'effort est incertain et dépend de la taille de chaque effet dans la relation d'agence.
Barkema (1995) confirme les intuitions proposées par Frey (1993) : plus le contrat psychologique est important dans la relation d'agence, plus la surveillance du principal dégrade l'effort de l'agent. Dickinson et Villeval (2008) ont affiné la théorie de l'éviction de la motivation liée à la surveillance en ajoutant que lorsque la surveillance était au-delà d'un certain seuil d'intensité, elle supplantait la motivation. De plus, lorsque le partage du résultat doit être négocié à travers un paiement variable de l'agent, la surveillance peut réduire la motivation à réaliser l'activité. Dans leur étude, ils confirment également l'intuition selon laquelle lorsque la relation est interpersonnelle, l'effet d'éviction domine l'effet disciplinant du comportement de l'agent. Enfin, Arnaud et Chandon (2014) concluent également que la surveillance peut réduire la motivation à réaliser un effort.

Cowen et Glazer (1996) montrent, dans un modèle avec une rémunération exogène fixe, que l'agent peut éviter l'effort attendu si le principal augmente la surveillance. Ils utilisent l'analogie de l'examen de connaissances : pour valider un examen reposant sur une question unique qui couvre l'ensemble d'un programme un étudiant devra fournir un effort important ; alors que s'il s'agit d'un examen qui repose sur plusieurs questions, l'étudiant devra fournir un effort moins important. Dans ce cadre, la surveillance correspond à la règle de l'examen de connaissances et elle induit une dispersion de l'effort, si le principal souhaite obtenir davantage d'informations sur l'apprentissage de l'étudiant.
Dans leur modélisation, à chaque période, l'agent fait le choix de fournir un effort positif ou d'éviter l'activité. Son utilité augmente avec la probabilité d'obtenir la récompense, mais elle décroit avec l'effort (elle augmente avec l'évitement de l'activité). S'il esquive l'activité, le principal perd du profit. Le principal reçoit alors des messages pour inférer le comportement de l'agent. Lorsque l'agent fournit un effort suffisant, le principal verse une rémunération fixe qui dépend de la probabilité de recevoir un volume de messages qui signalent l'effort réalisé ou un effort évité. Enfin, la probabilité de recevoir la récompense dépend positivement du niveau d'effort.
En raisonnant en information imparfaite, une augmentation du nombre de messages reçus sur le comportement de l'agent peut induire une diminution de l'effort. Le mécanisme est le suivant : une augmentation du nombre de messages reçus par le principal provoque un effet revenu qui augmente la probabilité de recevoir la récompense et si l'évitement est un bien normal, alors cela incite à davantage éviter l'effort. Cette augmentation provoque un effet substitution vers plus ou moins d'évitement. Par conséquent, restreindre le nombre de messages reçus peut être un meilleur dispositif d'incitation.

Les auteurs interprètent cette situation par le phénomène suivant : lorsque le principal reçoit des informations qui montrent que l'agent fournit l'effort qui répond à leurs attentes pendant une certaine période, alors ce dernier peut relâcher son effort à la période suivante, puisqu'il est certain d'obtenir la récompense ; le mécanisme peut également se transposer à une situation où l'agent qui est détecté comme étant défaillant voit la probabilité d'être récompensé lui échapper, ce qui l'incite à continuer à ne pas fournir d'effort attendu.

Ichino et Muehlheusser (2008) montrent que la surveillance forte d'un partenaire au début de la relation n'est pas optimale si l'objectif est de déterminer sa loyauté et si le coût de rupture de la relation s'accroît avec le temps. L'intuition est la suivante : en surveillant trop, on apprend moins sur le comportement du partenaire hors surveillance. C'est seulement en donnant du temps au partenaire la possibilité de mal se conduire, s'il était tenté par cette attitude, et seulement dans ce cas, qu'il y ait une chance de prendre connaissance de son comportement à un moment où la séparation est possible à moindre coût.

Enfin, Falk et Kosfeld (2004) en expérimentant la possibilité pour le principal de modifier le niveau de surveillance de l'agent, par l'implémentation d'un critère minimal de performance avant que l'agent ait choisi un niveau d'effort, montrent que la surveillance entraine des coûts cachés, notamment à cause de la réduction de l'effort des agents. Lorsque les individus réagissent négativement à la surveillance, ils la perçoivent comme un signe de méfiance et une limitation de leur autonomie.

Dans ce cadre, le concept de surveillance est précisé dans le sens où le principal restreint le domaine de choix de l'agent, notamment les choix opportunistes. Lorsque le domaine de choix est étréci, la majorité des agents choisit un plus faible niveau de performance par rapport à une situation où le principal leur fait confiance. La variabilité du domaine de choix est non-monotone, pour des seuils minimaux faibles de performance et la surveillance induit relativement moins d'effort que dans une situation de confiance, tandis que pour des seuils minimaux élevés de performance, la surveillance entraine autant d'effort que dans le cas d'une surveillance nulle.

Les travaux issus de cette recherche ont précisé l'apparition des coûts cachés. Ziegelmeyer et al. (2012) expliquent que les coûts cachés apparaissent lorsque les compensations sont hypothétiques. Schnedler et Vadovic (2011) invoquent la notion de légitimité de la surveillance. Von Siemens (2013) invoque quant à lui l'idée de la réciprocité.

III. La problématique générale : identifier, modifier, améliorer le rôle des dispositifs de vérification de l'information

À l'issue de cette présentation de la littérature issue de l'économie des dispositifs de vérification de l'information dans les organisations, une question majeure se pose : de quoi s'agit-il ? En effet, bien que nous ayons montré qu'il existait des schémas contractuels optimaux destinés à faire aligner les intérêts divergents des individus de l'organisation pour assurer la coopération entre les parties, et que ces derniers avaient des effets disciplinant ou des effets d'éviction, force est de constater que de telles architectures semblent véritablement complexes à appliquer dans la pratique. Les tenants de l'économie des contrats jugent eux-mêmes irréaliste l'implémentation de ces dispositifs.

Pourtant, évoquer la surveillance semble aller de soi : c'est une pratique qui consiste à vérifier ce qu'il se produit, à être certain que ce que nous observons est exactement conforme à ce que nous constatons. C'est une pratique qui semble communément répandue et dans le même temps, il ne semble pas y avoir dans la littérature une véritable disposition à éclaircir ce que revêtent concrètement ces dispositifs de surveillance. Quels sont leurs paramètres ? Quelle est leur étendue ? Quelles sont les grandeurs pertinentes à mesurer ? Où se situent-ils ? À quel moment sont-ils appliqués ? Qui les utilise ?

Selon nous, il existe des marges de manœuvre pour identifier plus rigoureusement les dispositifs de vérification de l'information. L'identification permet de progresser sur l'ingénierie de la surveillance et d'être capable d'embrasser l'ambivalence de son utilisation sur les comportements individuels.

La matrice de la surveillance dépend du cadre organisationnel dans lequel elle se positionne. Or, la théorie des contrats complets s'est essentiellement consacrée à analyser la relation d'emploi employeur-employé, la relation de gouvernance, propriétaire-directeur, et la relation d'assurance. Pourtant, la relation d'agence ne se borne pas à ces trois cas de figure. Il est souhaitable de pouvoir faire avancer la connaissance pour approfondir ce cadre d'analyse et de l'enrichir grâce à des investigations dans d'autres organisations productives. À partir des caractéristiques de la relation d'agence préalablement définies qui constituent un faisceau d'indices, nous postulons qu'il est possible de distinguer des relations alternatives ressemblant à celles traditionnellement analysées.

Pour répondre à cette première partie de la problématique, nous avons fait le choix de nous positionner sur le terrain et de mener un travail empirique. La confrontation du chercheur à l'épreuve des faits et du contexte social et culturel est une condition *sine qua non* de la

compréhension des dynamiques et des processus qui environnent un phénomène et des pratiques.

Une fois que le cadre organisationnel est défini, une série de questionnements se pose : que se passe-t-il quand la configuration des paramètres du dispositif de vérification de l'information est modifiée ? Dans quelle mesure une configuration particulière du dispositif de vérification de l'information peut-elle s'appliquer à des contextes organisationnels différents ? Quels sont les gains et les bénéfices engendrés par un système particulier ?
Pour répondre à ces questions, il importe de tester les effets des différents paramètres des dispositifs de la vérification de l'information sur les comportements individuels, sur des variables d'informations dont le principal et l'agent ne disposent pas dans les mêmes proportions.
L'objectif de la variation d'un dispositif est de résoudre un problème d'efficacité dans le modèle de répartition des ressources. En l'occurrence, dans le cadre organisationnel de la relation d'agence, le problème primordial est un problème de coopération. Par conséquent, il est nécessaire de savoir si la modification du système de surveillance permet d'améliorer la coopération en répartissant différemment des ressources pour atteindre cet objectif.
Enfin, être capable d'estimer les effets d'une modification, c'est finalement être capable d'évaluer un lien de causalité. Dans quelle mesure ce que nous observons peut-il correctement être inféré du changement qui a été subi et indépendant d'autres considérations sociales ou environnementales ?

Ces questions nous ont conduits à adopter la méthode de l'expérimentation contrôlée de terrain par assignation aléatoire (« *randomized controlled trial* »), afin d'étudier le rôle que jouent les paramètres des dispositifs de vérification de l'information sur les comportements des parties prenantes à des relations potentiellement similaires à une relation d'agence, en les modifiant et en évaluant l'effet de ces modifications.

IV. L'expérimentation contrôlée de terrain avec assignation aléatoire

L'essentiel du travail empirique en économie consiste à pouvoir identifier un lien de causalité entre une variable exogène indépendante d'un terme d'erreur et une variable dépendante (Kramarz et al., 2006). Dans ce cadre, deux problèmes récurrents se posent. D'une part la causalité est contrainte par la question de la sélection, qui suppose d'être certain que ce qui est observé sur les variations de la variable exogène est indépendant d'autres considérations, ce qui renvoie à la fameuse maxime sous-jacente *ceteris paribus*, « toute chose étant égale par ailleurs ». D'autre part, la causalité est conditionnée par la question de l'évaluation, qui exige d'être capable

de tenir compte des fluctuations de la variable si un évènement n'avait pas lieu, ce qui renvoie à la définition du contrefactuel, c'est-à-dire la réflexion sur les évènements qui ne se sont pas réalisés, mais avait pu advenir sous certaines conditions.

L'économétrie est la discipline qui se propose de répondre à ces deux problèmes en construisant des modèles de vérification des hypothèses de relation et en utilisant les données issues de l'observation statistique des variables pertinentes pour estimer des paramètres de la théorie économique. La critique de Leamer (1983) a provoqué une remise en question de la crédibilité des preuves issues de l'économétrie, notamment en insistant sur l'idée que « presque personne ne prend au sérieux l'analyse des données et plus sûrement, presque personne ne prend au sérieux l'analyse des données réalisées par les autres ». Il est sous-entendu que lorsque les spécifications retenues ou la forme fonctionnelle des modèles économétriques changent, alors les résultats obtenus sont également modifiés. Dans sa critique, Leamer (1983) proposait de recourir aux tests de sensibilité. Bien que cette méthode fût salutaire pour la discipline parce qu'ils permettent de renforcer la robustesse des résultats obtenus, ces tests n'ont en revanche pas été révolutionnaires pour la méthode économétrique.

La révolution de crédibilité selon Angrist et Pischke (2010) est le recours à des protocoles expérimentaux qui sont associés, dans une certaine mesure, à l'assignation aléatoire d'une modification de l'environnement d'action. L'origine de cette révolution méthodologique pour la science économique démarrerait à partir des travaux de Lalonde (1986), qui a comparé les résultats d'une évaluation économétrique d'un programme social avec ceux d'un essai avec assignation aléatoire et conclut que cette dernière prodigue des résultats plus robustes, et ceux d'Ashenfelter (1987), qui concluent que l'innovation dans les spécifications alternatives est toujours devancée par l'assignation aléatoire. Par conséquent, l'expérimentation par assignation aléatoire serait la seule méthode crédible. Elle devrait ainsi être le « *gold standard* » de l'économiste appliqué.

Néanmoins, il est important de rappeler que l'expérimentation n'est pas quelque chose de fondamentalement révolutionnaire. Chacun expérimente, c'est-à-dire qu'on acquiert des connaissances par la pratique, en se soumettant à un ensemble de processus qui contribue à atteindre un résultat précis. Dans cette notion est sous-entendue l'idée que l'expérimentateur détient un plus ou moins grand degré de contrôle sur le changement d'environnement (Harrison, 2013). Ainsi, une expérimentation « scientifique » sera décrite comme « contrôlée » lorsque l'investigateur manipule l'environnement pour introduire un changement pour mesurer les conséquences de ce changement sur le sujet sous étude. Lorsque le changement est indépendant de la volonté d'un expérimentateur ou qu'il ne fait pas l'objet d'une volonté de modifications a priori, on parlera d'expérimentation « naturelle ».

Le contrôle de l'expérimentateur est un premier moyen de répondre au problème de sélection, car il permet de définir la stratégie d'implémentation en donnant les critères d'encadrement de l'environnement dans lequel s'étudie le lien de causalité, les conditions de recrutement des sujets, les modalités d'assignation des sujets au changement analysé et les procédures de collecte des données de qualité (Bloom, 2008). En revanche, le problème de l'évaluation demeure pour trouver des mesures fiables de l'effet causal, la base implicite de comparaison étant l'état primitif du sujet sous étude avant que le changement soit délivré. Cet état n'est jamais observable pour un même sujet, parce qu'il est impossible de retourner dans le passé (Rubin, 1974).

Nous devons à Fisher (1935), l'intuition selon laquelle la méthode la plus satisfaisante pour atteindre l'équivalence entre un sujet qui bénéficie du changement et un sujet témoin, qui n'en bénéficie pas, est de s'assurer que les sujets sont assignés au changement de manière totalement aléatoire. La différence constatée sur les résultats moyens des sujets qui ont subi le changement et ceux qui n'ont pas subi le changement après l'incorporation du changement peut être directement attribuée à l'effet causal moyen du traitement.

Pour montrer cette relation, prenons une situation hypothétique, avec un individu i qui est soit exposé à un traitement T soit qui ne le subit pas C (Rubin, 1974). Il existe une phase d'initialisation de la situation avec traitement et sans traitement et une phase d'observation du résultat. Le traitement est assigné aléatoirement, c'est-à-dire qu'il existe un mécanisme qui attribue à l'individu la même probabilité de recevoir le traitement. L'objectif est de déterminer l'effet causal « typique » du traitement par rapport à une situation sans traitement sur une variable d'intérêt Y. On note $Y_i(T)$ la valeur du résultat étant donné que l'individu a reçu le traitement et $Y_i(C)$ la valeur du résultat étant donné que l'individu n'a pas reçu le traitement.

L'effet causal pour un individu particulier sur période donnée est la différence entre ce qu'il serait arrivé pendant la période d'observation si l'unité a été exposée au traitement et ce qu'il serait arrivé pendant la période d'observation si l'unité n'a pas été exposée au traitement. C'est la différence de résultat $Y_i\ (E) - Y_i\ (C)$. Or, il est impossible de mesurer en même temps les deux résultats puisqu'il est impossible de retourner dans le passé pour donner le traitement ou pour ne pas donner le traitement.

En généralisant à I individus, on peut écrire que l'effet causal est la moyenne des effets causaux individuels, ce qui correspond à :

$$\Delta = \frac{1}{I} . \sum_{i=1}^{I} \big(Y_i(T) - Y_i(C)\big)$$

Par conséquent, l'effet « typique » de causalité est l'effet causal moyen pour I individus. En revanche, il est toujours impossible d'observer en même temps pour un même individu son résultat sous traitement et son résultat sans traitement.

Une première étape du raisonnement pour estimer l'effet causal est de rassembler chaque observation individuelle dans un groupe traité, c'est-à-dire tous ceux qui ont bénéficié du traitement et dans un groupe témoin, c'est-à-dire tous ceux qui n'ont pas bénéficié du traitement. On peut donc écrire :

$$\Delta = (E(Y(T) \,|\, T = 1) - E(Y(C) \,|\, T = 1))$$

où le premier terme de la différence correspond à l'espérance mathématique du groupe traité lorsqu'il reçoit le traitement qui est observable et le second terme correspond à l'espérance mathématique du groupe traité lorsqu'il ne reçoit pas le traitement qui est inobservable.
Une deuxième étape du raisonnement est de faire l'hypothèse que l'un des deux groupes est le contrefactuel de l'autre, c'est-à-dire qu'ils se ressemblent et que le groupe non traité évolue en l'absence de traitement de la même manière que l'aurait fait le groupe traité en l'absence de traitement. Cela correspond à :

$$E(Y(C)|C = 1) = E(Y(C) \,|\, T = 1)$$

Ainsi, on peut écrire :

$$\Delta = (E(Y(T) \,|\, T = 1) - E(Y(C)|C = 1))$$

Et faire apparaître le contrefactuel :

$$\Delta = (E(Y(T) \,|\, T = 1) - E(Y(C) \,|\, T = 1) + E(Y(C) \,|\, T = 1) - E(Y(C)|C = 1))$$

Dans cette expression, on distingue à gauche l'effet causal et à droite la différence entre le résultat moyen qu'auraient obtenu ceux qui sont traités s'ils ne l'avaient pas été et le résultat de ceux qui ne sont pas traités. Cette quantité représente le biais de sélection.

L'assignation aléatoire permet de vérifier l'hypothèse qu'en moyenne les deux groupes sont similaires et que leur résultat moyen en l'absence de traitement est identique. En effet, ce processus de tirage au sort permet d'enlever la corrélation systématique entre le statut du changement et les caractéristiques observables et inobservables des sujets sous étude, c'est-à-dire qu'il permet d'enlever définitivement le biais de sélection. C'est donc un moyen supplémentaire de contrôler le changement étudié.
L'estimateur de la différence entre les deux variables de résultats est un estimateur sans biais de l'effet causal (Rubin, 1974). Par conséquent, il est possible de se prémunir contre le problème de validité interne des résultats estimés parce qu'il rend aléatoire toute source d'incertitude à propos de l'estimation de l'impact pour un échantillon donné.

Il est toujours possible du fait même de sa nature aléatoire que le tirage au sort d'un échantillon entraine la constitution d'un groupe traité porteur de caractéristiques préexistantes différentes de celles du groupe témoin. Cela a pour conséquence de surestimer ou de sous-estimer l'effet moyen du traitement. Cependant, plus la taille des groupes est importante, moins ce risque peut apparaître.
Enfin, la simplicité de la méthode rend plus convaincante la production des résultats par rapport à des techniques d'estimation innovantes, mais potentiellement difficilement interprétables compte tenu des différentes hypothèses de mise en application (Burtless, 1995).

L'expérimentation contrôlée par assignation aléatoire en sciences sociales a évolué au fur et à mesure des époques et des contextes de mise en œuvre ce qui a engendré un large spectre de méthodologies. Dans ce spectre, il existe un arbitrage entre le contrôle et le réalisme de l'environnement dans lequel se réalise le changement étudié (List et Reiley, 2008).
D'un côté du spectre se situent les expérimentations en laboratoire, où l'environnement est rendu artificiellement « stérile » grâce aux instructions de l'expérimentateur qui permettent d'isoler les effets d'un traitement exogène sur le comportement d'individus. Dans le laboratoire, l'expérimentateur impose toutes les hypothèses structurelles de modélisation d'une théorie (préférences induites, les institutions d'échanges, l'ordre de mouvement dans le « jeu ») et examine comment les joueurs se comportent par rapport à la théorie.
De l'autre côté du spectre se situent les expérimentations de terrain qui utilisent un contexte d'étude réel et qui étudient le comportement des individus sans induire d'instructions. Sur le terrain, l'expérimentateur admet que les préférences et les institutions correspondent à ce qui se produit dans le monde réel en testant les hypothèses structurelles et les hypothèses comportementales.

Harrisson et List (2004) ont identifié cinq éléments complémentaires qui permettent d'affiner la séparation entre le laboratoire et le terrain afin de proposer une taxonomie de référence de l'expérimentation contrôlée par assignation aléatoire.
Tout d'abord, la nature de l'échantillon permet de distinguer le type d'expérimentation. Fréquemment, les individus qui participent aux essais en laboratoire sont des étudiants, puisque les chercheurs travaillent principalement sur les campus universitaires. Le recours à des sujets « de terrain » permet d'identifier des échantillons où les variables sociodémographiques (âge, niveau d'études et pratiques sociales) sont plus ou moins concentrées par rapport à une population étudiante. Alternativement, l'échantillon peut représenter une population cible de l'économie ou la population générale.

Sur le terrain, les sujets ont certaines informations quant à leurs activités d'échange en plus de leurs connaissances sur les institutions d'échange. Dans un environnement stérile, l'importance de ces éléments est diminuée par construction et cela peut conduire à des changements comportementaux.

La nature du bien échangé, qu'il soit réel ou artificiellement défini, est une autre façon d'identifier les deux approches. Cette dichotomie est poreuse puisque des biens physiques sont utilisés dans le cadre du laboratoire (Ruström, 1998). Davantage que la nature du bien, c'est la manière dont sont provoquées les estimations des préférences qui permet de caractériser une expérimentation de terrain ou de laboratoire.

La valeur accordée aux données d'un problème spécifique est également un critère de choix important qui conditionne l'environnement de l'expérimentation. Dans un environnement réel, l'expérimentateur laisse la possibilité aux sujets de posséder plus ou moins de degrés d'heuristique, c'est-à-dire la capacité d'apprentissage et de résolution des problèmes pour des tâches spécifiques. Une question importante est soulevée par cette dimension : est-ce que le succès à résoudre un problème donné dans un contexte réel peut se transposer dans un autre contexte qu'il soit abstrait ou réel ?

Enfin, la nature des enjeux permet d'identifier le type de protocole expérimental. Les enjeux du laboratoire sont souvent moins importants, par exemple lorsque les valeurs monétaires des biens sont converties pour réduire le coût de l'intervention (100 € = 1 unité de compte expérimentale). Le calcul coût-avantage peut alors devenir imprécis.

Cette séparation n'est pas aussi stricte qu'elle pourrait le faire penser et de nombreuses expérimentations de terrain font un détour par le laboratoire pour confirmer des résultats et réciproquement (Levitt et List, 2009).

Avant que la méthode expérimentale contrôlée de terrain par assignation aléatoire soit conçue comme la seule méthode empirique « crédible » aux yeux d'un certain nombre d'économistes (Angrist et Pischke, 2010), sa pratique a été mise en valeur par différents courants de recherche (Levitt et List, 2009 ; Monnier, 1992). Les économistes l'ont déployée relativement récemment par rapport à d'autres sciences sociales comme la psychologie où la tradition expérimentale est plus ancienne. Ces courants ont en commun d'avoir subi plusieurs phases d'émulation intensive suivie de relative déception (Labrousse, 2010).

Le programme de recherche qui a le plus appliqué cette méthode correspond à celui de l'évaluation d'impact des politiques publiques que l'on catégorise plus souvent sous le terme d'expérimentations sociales (Heckman et Smith, 1995 ; L'Horty et Petit, 2011 ; Ferber et Hirsch, 1982). La plupart du temps, il s'agit d'une étude financée par des fonds publics qui incorpore un

protocole expérimental et dont les aspects expérimentaux sont appliqués sur une période donnée, sur un ou plusieurs segments de population, et dont l'objectif est d'évaluer les effets sociaux et économiques agrégés d'un dispositif public expérimental. D'un point de vue historique, cette caractérisation de l'expérimentation contrôlée de terrain avec assignation aléatoire correspond à la vague de programmes sociaux des années 1960 et 1970 aux États-Unis pour combattre la pauvreté et comprendre les élasticités prix et revenu de la demande des ménages, ainsi qu'à la vague de programmes d'aide des pays en voie de développement à la fin des années 1990 (Barnejee et al., 2015) et à la vague de rationalisation des dépenses publiques dans les pays européens au milieu des années 2000. Dans une certaine mesure, les campagnes de tests par correspondance pour identifier et mesurer les discriminations sur le marché du travail sont également un avatar des expérimentations sociales.

Un deuxième champ de recherche entreprend de développer des protocoles expérimentaux appliqués à un contexte réel sur des sujets qui ne savent pas qu'ils font partie d'une expérimentation (Harrisson et List, 2004). Ces expérimentations de terrains se sont davantage construites par opportunité et à une échelle inférieure à celles des expérimentations sociales. Elles ont traditionnellement plus d'ambition théorique que les expérimentations sociales, notamment parce qu'elles transposent davantage de protocoles de recherche en laboratoire sur le terrain et elles affichent davantage leur volonté d'estimer les paramètres de la théorie économique. Selon ses promoteurs, elles produisent des résultats qui permettent de faire de plus fortes inférences que celles permises dans le laboratoire sur des données non contrôlées. Ces expérimentations « artéfactuelles » aident à déterminer si les études en laboratoire ou sur le terrain doivent être réinterprétées ou définies plus précisément. Enfin, elles permettent d'aider à déterminer les conditions nécessaires à la mesure d'un lien de causalité en fonction de type de données observées sur le terrain ou dans le laboratoire.

Un troisième courant de recherche, moins apparent dans la littérature, mais qui refait surface depuis le début des années 2010, coïncide avec la phase d'approfondissement paradigmatique de l'économie des organisations (Bandiera et al., 2011 ; Shaw, 2014). Plutôt laissées de côté par les économistes, ces expérimentations dans le cadre des organisations approfondissent les questions théoriques traditionnelles : la rémunération des employés, l'organisation du travail par équipe, la fixation des prix, les caractéristiques des produits, la technologie de production, le management, etc. Ce mouvement est impulsé d'une part par le déploiement d'investissements des organisations pour améliorer leurs performances en termes de productivité, et d'autre part par la volonté affichée par les chercheurs d'investiguer la boîte noire de l'organisation. Les expérimentations traitant des phénomènes se réalisant à l'intérieur des organisations sont généralement conçues pour mettre en lumière les dispositifs de résolution des problèmes d'agence (Bandiera et al., 2011).

Bien que ces expérimentations aient connu une résurgence récente, elles sont en réalité anciennes. En effet, la première série d'expérimentations de ce type a été conduite dans l'établissement Hawthorne de l'entreprise de production d'électricité Western Electric. Ces expérimentations, menées entre 1924 et 1927, avaient notamment pour but d'identifier le rôle de la variation de l'intensité lumineuse sur la performance des ouvrières de l'établissement (Mayo, 1933). Une autre expérimentation influente a étendu la portée des expérimentations dans les organisations en étudiant les pratiques innovantes de travail sur la productivité dans des ateliers de vernissage d'acier (Ichniowski et al., 1997).

Nous nous inscrivons naturellement à la suite de ce troisième courant d'expérimentations traitant des phénomènes internes à l'organisation productive.

V. Les précautions à envisager dans cette recherche et la présentation des chapitres

Toutes les expérimentations de terrain permettent de collecter des données (List et Reiley, 2010). La collecte des données permet de construire une théorie, de tester les prédictions de la théorie et de mesurer des paramètres clés en assumant que la théorie est correcte.

De plus, les protocoles expérimentaux s'ils ont correctement décrits ont la capacité d'être répliqués, c'est-à-dire qu'il existe la possibilité de reproduire l'expérimentation pour tester si les hypothèses peuvent être vérifiées ou non dans d'autres conditions (Levitt et List, 2009). Ils permettent d'induire la génération de données comparables. La réplication peut s'effectuer soit, en réanalysant les données d'une expérimentation afin d'obtenir les mêmes résultats, soit en réalisant une expérimentation alternative avec le même protocole pour mesurer des résultats similaires ou enfin en testant les hypothèses avec un autre protocole.

Cependant, les expérimentations de terrain contiennent plusieurs limites qui ont sans doute compromis leur développement et qui témoignent de la relative déception des chercheurs dans chaque courant à l'issue de leur utilisation (Labrousse, 2010 ; Jatteau, 2013).

Premièrement, il existe des limites éthiques, car il s'agit le plus souvent d'expérimenter un dispositif contrôlé qui contient une part d'incertitude sur les effets produits sur une population humaine. D'une part, le traitement peut faire supporter des risques importants aux individus traités, et d'autre part, il peut être intolérable d'empêcher les individus du groupe sans traitement de ne pas bénéficier d'un traitement aux effets positifs. La question que pose Fougère (2012) est de savoir « s'il est plus juste ou plus bénéfique de généraliser, éventuellement en le rendant obligatoire, un programme qui risque d'être préjudiciable, de ne pas rendre accessible à tout un chacun un programme qui peut s'avérer profitable ou bien de mettre en place préalablement à toute décision de généralisation ou d'abandon définitif du programme une procédure d'évaluation

par expérimentation avec l'aide et le consentement éclairé d'un nombre limité d'individus dont certains acceptent de se voir refuser à l'issue du tirage au sort et à l'exclusion de toute considération le droit d'accès au programme le temps du dispositif ? »
Les participants à l'expérimentation doivent ainsi être informés des gains et pertes engendrés par le traitement et de recevoir une compensation en cas de tort subi : c'est la notion de consentement éclairé. Dans ce cadre, il est légitime d'appliquer un principe de précaution lorsque l'expérimentateur ne sait pas quels sont les effets mesurés.

Deuxièmement, il existe des limites logistiques, lorsque la conception et l'implantation d'une expérimentation sont coûteuses en ressources. Elles consomment des ressources financières importantes : expérimentateurs, matériel expérimental, compensations aux participants à l'expérimentation.
Elles consomment également du temps qui pourrait être consacré à d'autres activités, notamment lorsque l'expérimentateur souhaite contrôler l'environnement du traitement pendant des périodes de perturbations qui n'ont pas été anticipées, par exemple si la conjoncture économique se détériore.
Elles sont coûteuses pour l'expérimentateur, mais également pour les parties prenantes qui sont directement ou indirectement concernées par l'implantation d'un dispositif. En effet, les personnes qui ne subissent pas le traitement, mais qui doivent l'appliquer, peuvent être opposées à certaines modifications, par exemple dans le cas d'une modification d'un dispositif d'accompagnement social. Cette résistance au changement peut réduire la portée du dispositif.

Troisièmement, l'expérimentation peut être contrainte par des biais d'échantillonnage lorsqu'il existe un problème de recrutement (faiblesse des effectifs), un problème d'attrition (certains participants quittent l'expérimentation avant la fin), un problème dans le processus d'assignation aléatoire des participants à l'expérimentation (les groupes différents de la population de référence) ou encore un problème de substitution (les participants cherchent à recevoir le traitement alors qu'ils sont affectés au groupe témoin et inversement).

Quatrièmement, il existe des limites d'inférence. Une expérimentation est toujours locale, ciblée et temporaire. Ces trois caractéristiques peuvent généralement empêcher d'étendre la portée des résultats dans un cadre plus général. Cela pose un problème de validité externe des résultats (Heckman, 1997 ; Deaton, 2010).
Une critique connexe est que l'expérimentation permet de donner de « bonnes » réponses sans de « bonnes » questions. L'idée sous-jacente est qu'il est délicat de distinguer les résultats entre deux théories alternatives lorsque le support théorique à l'expérimentation est difficilement perceptible.

De plus, les individus étant sous une condition expérimentale modifient leur comportement précisément parce qu'ils y participent, et cela peut se traduire par une augmentation de la désirabilité sociale liée à la gratitude de faire partie d'un projet de recherche. C'est l'effet Hawthorne : dans l'expérimentation de Mayo (1933) les chercheurs ont ainsi mesuré que la variation de la lumière avait un effet positif sur la performance, quelle que soit la variation de l'intensité lumineuse, ce qui a montré que, ce serait davantage l'intérêt porté au travail des ouvrières qui aurait contribué à augmenter leur performance, que l'effet direct de la lumière.

Il s'agit alors pour l'expérimentateur de tenir compte de l'ensemble de ces limites pour proposer des protocoles suffisamment rigoureux afin de modifier l'environnement pour identifier sans biais les liens de causalité. Dans la mesure du possible, les protocoles que nous avons utilisés dans notre thèse de doctorat ont été conçus pour être à même de contenir ces limites. Néanmoins, il reste très ambitieux de vouloir prétendre à un contrôle total de la condition expérimentale ainsi qu'à la production de lien de causalité directement mesurable. Nous appliquerons dans ce contexte un principe de précaution pour interpréter nos résultats.

Cette thèse de doctorat se situe donc au confluent de plusieurs champs de recherche et émerge d'un constat partagé par la discipline d'un approfondissement des connaissances sur les mécanismes de l'organisation quant à l'implantation de dispositifs de vérification de l'information dans le cadre de relation d'agence où l'aléa moral est susceptible d'apparaître. Nous recourrons à l'expérimentation de terrain par assignation aléatoire pour ouvrir la boîte noire de la surveillance et apprécier l'économie des dispositifs de vérification de l'information.

Dans notre premier chapitre, nous avons envoyé des messages ciblés à des entreprises du secteur marchand contenant des informations sur le pouvoir de vérification et de sanctions de l'Agence centrale des organismes de sécurité sociale (Acoss) afin de diminuer la fraude sociale. Nous avons ainsi testé l'effet de la variation de la quantité d'informations détenues sur le dispositif de vérification de la déclaration sociale des entreprises sur des indicateurs de masse salariale en stock et en flux.
L'Acoss fait face à un aléa de consentement à la contribution sociale de la part des cotisants. Lorsque cet aléa se réalise, il se traduit comme de la fraude aux prélèvements, ce qui réduit les sources de financement de la protection sociale. L'agence de recouvrement dispose d'un outil de contrôle comptable des cotisants ainsi que d'inspecteurs chargés d'investiguer les entreprises aléatoirement ou sur la base d'un faisceau d'indice.

Le dispositif communicationnel expérimental complète les mécanismes de contrôle et d'incitations au service de l'administration pour contrer les tentatives de dissimulation de la main-d'œuvre et les comportements de fausses déclarations de rémunération. Les messages contenaient des informations sur les paramètres canoniques des modèles de fraude aux prélèvements fiscaux ainsi que des renseignements sur des paramètres enrichis d'interactions sociales. Nous montrons que plusieurs messages d'informations ont augmenté la déclaration d'embauche de personnes en contrat à durée indéterminée tandis que des établissements réduisaient leurs recrutements en contrat à durée déterminée. Nous interprétons ce phénomène comme un encouragement à améliorer la déclaration sociale qui sécurise les relations de travail durables. Nous contribuons ainsi à réduire l'aléa moral de la déclaration sociale et à proposer une alternative aux dispositifs coercitifs.

Dans notre deuxième chapitre, nous avons modifié l'intensité de la surveillance électronique des conseillers d'un centre d'appels sous-traitants afin d'améliorer les conditions de travail sans perturber l'organisation du travail. Nous avons ainsi testé l'effet de la variation de quantité d'information collectée par le dispositif de surveillance sur l'effort produit par les salariés de l'entreprise à partir de la mesure de la performance individuelle et d'indicateurs de la qualité de vie au travail.

Le cadre du centre d'appels permet d'identifier la relation d'emploi et le dispositif de surveillance de la productivité des salariés. Les centres d'appels sont associés à l'image d'une usine moderne, où l'activité est stéréotypée. La surveillance du travail est paramétrée pour restreindre l'opportunisme des employés qui ne fourniraient pas l'effort demandé et pour permettre d'appliquer un système de primes salariales.

Ce chapitre est une contribution à la littérature émergente qui tente d'analyser les conséquences d'un environnement d'évaluation et de contrôle permanent de la productivité des employés sur leur qualité de vie au travail et sur leur performance. Pendant seize semaines, quarante-six conseillers ont participé à cette étude qui consistait à affecter aléatoirement un groupe d'employés à la condition expérimentale, qui diminuait l'intensité de la surveillance électronique et un autre groupe d'employés dans une condition témoin. L'intensité de la surveillance électronique de la performance a été réduite grâce à la suppression de l'accès à des indicateurs en temps réel à la fois sur les consoles des superviseurs du plateau téléphonique, mais également sur les applications de suivi des appels des conseillers téléphoniques. Contrairement à nos hypothèses, une absence d'impact significatif sur la performance et une amélioration de la qualité de vie au travail, nous avons trouvé que la performance a été améliorée sur les campagnes d'appels sortants, et que la qualité de vie au travail a été détériorée sous la condition expérimentale. Ces résultats peuvent être expliqués par la déstabilisation des conseillers téléphoniques, qui ont perdu l'accès à leurs

indicateurs de performance et par un ajustement comportemental des superviseurs qui ont tenté de maintenir un niveau de surveillance en utilisant des pratiques plus explicites de surveillance.

Enfin, dans notre troisième chapitre, nous avons impliqué des étudiants de licence dans la co-construction de la règle d'évaluation de leur participation pendant les travaux dirigés afin d'améliorer leur réussite à l'université. Nous avons ainsi testé l'effet de la négociation contractuelle du dispositif de vérification de la participation à partir de la mesure des notes aux examens et de leur probabilité de valider leur semestre.

Le choix de ce terrain de recherche permet d'étendre la portée de la relation d'agence à la relation d'instruction. L'enseignant utilise l'examen des connaissances pour certifier le niveau d'acquisition de l'étudiant et inciter les étudiants à réaliser un effort d'apprentissage. Cependant, à l'université, les règles d'examen n'ont pas été adaptées à un contexte de démocratisation de l'accès à l'enseignement supérieur et d'accueil de bacheliers de plus en plus nombreux. De plus, l'université fait face à un problème régulier d'échec des étudiants, notamment en première année de licence. Dans ce cadre, nous envisageons l'innovation dans les règles d'évaluation des étudiants comme un élément d'amélioration de leur réussite dans leur parcours de formation.

Nous partons ainsi du postulat que dans le cadre de la relation enseignant-étudiant, le dispositif de vérification de l'effort d'acquisition de connaissances et de compétences peut être négocié afin de parvenir à une situation socialement optimale, en réallouant plus efficacement les ressources de l'éducation. Nous montrons que ce dispositif expérimental produit un impact positif sur la réussite à l'université dans certaines conditions, notamment en fonction de la manière dont est exécuté le contrat d'évaluation négocié.

Chapitre 1.
Expliquer les conditions du dispositif de contrôle de la fraude sociale pour améliorer le consentement à la contribution sociale[4]

[4] Ce chapitre résulte d'un projet de recherche intitulé « Recherche Expérimentale pour Analyser les Causes du Travail Illégal et de la Fraude Sociale » (REACTIFS) coréalisé par Bruno Crépon, Cyril Hagneré, Yannick L'Horty et Pascale Petit et provient d'un partenariat entre la Fédération de Recherche (n°3435) Travail Emploi et Politiques Publiques (TEPP), le Centre de Recherche en Économie et Statistique (CREST) et l'Agence centrale des organismes de sécurité sociale (Acoss). Les premiers résultats de ce travail ont été présentés au cours de la Cinquième Journée annuelle des doctorants de l'ERUDITE (Jade) en juin 2017 et pendant la XIV^{ème} Conférence Annuelle sur l'Évaluation des Politiques Publiques à l'Université d'Angers en octobre 2017.

I. Introduction

En avril 2016, l'Agence centrale des organismes de sécurité sociale (Acoss), l'organisme français chargé du recouvrement de l'ensemble des contributions sociales destinées à financer le régime général de la Sécurité sociale, a publié une note d'étude et de résultats portant sur l'estimation de la fraude aux prélèvements sociaux. Selon cette évaluation « imprécise » de l'ampleur des irrégularités des entreprises soumises aux cotisations sociales, chaque année, entre 6,1 et 7,4 milliards d'euros seraient non recouvrés par l'organisme (Acoss, 2014), soit 1,5 % de l'ensemble des recettes issues des cotisations (Gonzales et al., 2018).

La fraude au prélèvement social se définit comme tout acte illégal consistant à réduire la perception obligatoire d'une partie du revenu destinée à financer un système de protection sociale. Celui-ci recouvre l'ensemble des mesures, dispositifs et prestations conçus pour diminuer, à l'échelle d'une collectivité humaine, les incertitudes de l'existence, notamment la pauvreté et la vulnérabilité, c'est-à-dire en réduisant l'exposition de ses bénéficiaires aux risques socialement identifiés et individuellement perçus (invalidité, accident, chômage, maladie, vieillesse et, plus récemment, perte d'autonomie) ainsi qu'en améliorant leur capacité à se protéger eux-mêmes contre les aléas et les pertes de revenu (adapté de Viet et du Bureau International du Travail). En France, le financement des prestations sociales s'effectue d'une part, sous la forme de prélèvements obligatoires réalisés par l'État et destinés à subvenir aux politiques publiques et d'autre part, sous la forme de cotisations, c'est-à-dire une contribution effectuée au nom d'une personne, par elle-même ou par autrui, afin d'acquérir ou de maintenir un droit futur à prestations.

Ces cotisations sociales représentent la majorité du financement de la protection sociale. En 2016, elles ont représenté 60,8 % des ressources totales. Il existe deux catégories de cotisations sociales : les cotisations effectives, c'est-à-dire les versements liés à l'emploi salarié ou indépendant aux organismes d'assurances sociales, et les cotisations imputées, c'est-à-dire les versements à un employeur pour des prestations spécifiques. Dans ce chapitre, nous nous concentrons exclusivement sur la fraude aux cotisations sociales effectives liées à l'emploi salarié qui représentent « trois quarts » des cotisations sociales (Gonzales, 2018). Ces cotisations effectives financent la quasi-intégralité des administrations de Sécurité sociale (ASSO) et sont recouvrées par l'Acoss.

Par conséquent, la fraude au prélèvement social impacte directement l'organisation du recouvrement, ses entités et sa gouvernance (Enoff et McKinnon, 2010) ainsi que la stabilité du financement de la protection sociale.

1. Les éléments de cadrage de la fraude sociale

Dans l'organisation de la protection sociale, c'est principalement le salarié qui finance la couverture sociale grâce à son salaire, c'est-à-dire la somme d'argent remis par l'employeur à l'employé en retour d'un travail exécuté. En effet, l'un des fondements de la Sécurité sociale contemporaine est de permettre de couvrir les frais engendrés par les risques de perte de la rémunération du travail. Le prélèvement repose sur l'idée de la déclaration, c'est-à-dire que la collecte s'effectue à partir du moment où les individus fournissent des renseignements sur l'existence des revenus.

Cependant, deux phénomènes concourent au déplacement de la responsabilité de la déclaration et du paiement des contributions sociales afin de rendre le recouvrement plus efficace. D'une part, l'entreprise tient une comptabilité de ses dépenses et de ses recettes, ce qui la rend plus à même de fournir des justifications des sommes versées à ses employés et de connaître sa masse salariale. D'autre part, le prélèvement des contributions sociales étant retenu lors de chaque paiement des salaires, il est plus efficace qu'un seul agent paye les salaires et les contributions sociales à partir de la même source. La conséquence de ces deux circonstances est que c'est l'employeur qui doit déclarer sa masse salariale et participer à la protection sociale à la place de ses employés.

Ainsi, l'entité économique qui réalise la fraude aux prélèvements sociaux est l'employeur et cela se traduit par des actions lui permettant de réduire l'incidence des prélèvements sur la masse salariale, c'est-à-dire la somme globale des salaires perçus par l'ensemble des employés d'une entreprise. En résumé, réduire l'incidence du prélèvement social revient, dans un système déclaratif, à ne pas déclarer l'intégralité du coût du travail. Ainsi, il existe deux types de comportements de mauvaise déclaration (Bailey et Turner, 2001) :

- lorsque la fraude porte sur les rémunérations, l'employeur déclare des rémunérations inférieures à celles versées aux salariés ou déclare des cotisations à payer inférieures à celles qu'il devrait verser à l'organisme de recouvrement. Il est également possible d'identifier ce type de fraude lorsque l'employeur règle en retard ou en partie des cotisations dues.
- En ce qui concerne la fraude qui porte sur l'effectif, l'employeur omet de déclarer un employé ou réalise de mauvaises déclarations pour affilier un employé à un mauvais régime, par exemple en déclarant qu'il est intérimaire ou stagiaire. Ce type de comportement correspond également au travail dissimulé, le délit qui consiste pour les entreprises à ne pas se conformer à leurs obligations légales de déclarations sociales et fiscales en termes d'activité et d'emploi salarié (Bouvier, 2007). Par conséquent, analyser la fraude aux cotisations sociales effectives revient également à analyser le recours au travail dissimulé, dans sa dimension sociale et d'emploi.

La fraude sociale est autant « le résultat d'opportunités », lorsque les entreprises réalisent de mauvaises déclarations ponctuellement, que « des actions dûment planifiées et orchestrées » où la fraude est organisée en réseau de manière à contourner sciemment la législation, voire à tirer profit des réglementations internationales (Oclti, 2009 ; Tian, 2011).

Il est raisonnable de penser que la fraude au prélèvement apparaît au même moment que se forme un système de prélèvement obligatoire (Torgler, 2008). Comme pour tout prélèvement, il existe un phénomène de non-consentement (Alm, 2018). Il est ainsi possible de dater l'apparition des premières formes de fraude sociale dès l'instauration de la Sécurité sociale en 1945, sans qu'il soit véritablement possible de vérifier cette assertion. La notion de non-consentement au prélèvement social peut également s'appliquer dans le contexte de la protection sociale (Goveia et Sosa, 2017). Ce n'est donc pas un phénomène nouveau.

Néanmoins, la fraude sociale a pris de l'importance à partir du moment où le législateur a souhaité réformer le financement de la protection sociale et améliorer l'efficacité du système (Rosanvallon, 1995 ; Barbier et Théret, 2009). Depuis la fondation de la Sécurité sociale, plusieurs phénomènes socio-économiques ont contribué à modifier son mode de fonctionnement. Nous pouvons évoquer les principaux événements majeurs : la montée du chômage liée au ralentissement de l'activité à la suite du choc pétrolier de 1973 ; l'élévation des dépenses de santé liée notamment au vieillissement de la population et à la meilleure prise en charge des soins grâce aux innovations médicales ; la croissance du nombre de départs en retraite liée au phénomène du baby-boom qui s'est transformé en papy-boom ; la prise en compte de nouveaux risques sociaux comme l'exclusion sociale ou la dépendance ; l'extension des couvertures sociales à de nouvelles catégories de population à l'instar des familles monoparentales. Ces phénomènes majeurs ont conduit à augmenter les dépenses de prestations sociales qui sont progressivement devenues supérieures aux recettes au tournant des années 2000 (Palier, 2005). Dans ce contexte, les organisations de sécurité sociale ont cherché de nouvelles sources de financement et la fraude sociale représente un vivier important de ressources à récolter. Ainsi, des comportements de dissimulation de la masse salariale et de non-consentement au prélèvement, qui étaient considérés comme tolérés et peu répréhensibles, sont finalement devenus des enjeux de politique publique pour récupérer une partie des ressources manquantes à la Sécurité sociale. Par conséquent, bien qu'étant un phénomène inhérent au prélèvement, la fraude est devenue récemment un enjeu majeur de la politique publique.

Au même titre que l'étude du non-consentement au prélèvement fiscal (Alm, 2012), les mesures de la prévalence de la fraude sociale sont rares et sujettes à certaines limites méthodologiques, l'information sur la non-déclaration étant, par définition, invisible (Joubert, 2009).

À la fin des années 2000, des estimations de la fraude sociale ont commencé à être produites pour alimenter l'idée selon laquelle la fraude était un phénomène à réduire afin d'augmenter les ressources de la Sécurité sociale. Ainsi, les premières estimations du Conseil des Prélèvements Obligatoires (2007) et de l'Acoss (2009) évaluaient des montants de fraude aux prélèvements sociaux compris entre 8 et 15,8 milliards d'euros. Ce manque à gagner aurait représenté 4 % du montant total des recettes de cotisations de l'époque.

La fraude sociale se concentre sur quelques secteurs de l'économie. Une étude réalisée à propos de deux mille entreprises à Shanghai révèle que ce sont dans les secteurs de la construction et de l'immobilier que la fraude sociale, entendue comme dissimulation du travail et réduction de la déclaration des salaires, était la plus importante. De plus, le fait d'avoir un grand effectif joue positivement sur la probabilité de frauder (Nyland et al., 2006). Les économistes de l'Acoss ont également estimé que la fraude sociale était la plus importante dans le secteur du commerce et l'hôtellerie-restauration (Acoss, 2014). Enfin, Joubert (2009) précise que les secteurs relevant de l'administration publique, de l'immobilier et les entreprises qui fournissent des services aux autres entreprises seraient des secteurs plus à risque que les autres. De façon générale, ce serait dans les services à forte intensité en main-d'œuvre que se développerait plus facilement le travail dissimulé, notamment parce que c'est au sein de ce secteur que se trouve la plus forte saisonnalité de l'activité et où la place de la sous-traitance permet des montages organisationnels destinés à la fraude.

2. Les conséquences sociales et économiques provoquées par la fraude sociale

Un système de protection sociale ne peut fonctionner qu'avec le support de ses participants (Rofman et Demarco, 1998). Par conséquent, la fraude au prélèvement social remet en cause le fonctionnement de la protection sociale, fragilise l'organisation du recouvrement et engendre des distorsions tant au niveau microéconomique des ménages et des entreprises qu'au niveau macroéconomique de la nation (Manchester, 1997).

La fraude sociale réduit les recettes de la protection sociale, ce qui diminue la pérennité financière de l'organisation. Cela restreint donc la capacité des organismes de protection sociale à piloter l'ensemble de leurs activités, de la collecte des contributions aux versements des prestations sociales. Par conséquent, elle limite l'adéquation entre la contribution et les droits sociaux des assurés, par exemple lorsque l'allongement des délais de versement de l'intégralité des cotisations augmente les délais de versement des prestations (McGillivray, 2001 ; Enoff et McKinnon, 2010). *In fine*, la fraude fragilise la légitimité et la crédibilité du système à fournir une protection optimale à ses assurés (McGillivray, 2001). Cette fragilisation du système peut engendrer plus de

comportements frauduleux, ce qui enclencherait un cercle vicieux, la fraude provoquant de la fraude.

Pour compenser la baisse des recettes engendrée par la fraude, l'organisme de sécurité sociale peut être contraint d'augmenter les taux de contribution au niveau individuel, ce qui accroît l'iniquité devant le prélèvement (Manchester, 1997). Cela peut se traduire par une réduction de la demande ou de l'offre sur les marchés.

La fraude provoque également un accroissement des inégalités de distribution de revenu pour des travailleurs similaires. En effet, dans un système contributif, la cotisation est un revenu différé. Par conséquent, réduire les contributions revient à restreindre les droits à prestations lorsque survient un risque social (maladie, vieillesse, chômage…). De plus, ne pas déclarer son effectif revient à nier le droit à la protection lié au contrat de travail qui encadre l'activité des salariés. La fraude provoque une distorsion de concurrence avec les contributeurs honnêtes, que l'on caractérise par du dumping social (Goveia et Sosa, 2017).

Enfin, au niveau macroéconomique, la fraude sociale provoque une distorsion au sein du marché du travail en faveur du secteur informel (Manchester, 1997). Ce dernier étant moins productif, la fraude peut réduire la croissance des richesses. Si l'on considère que la contribution sociale est de l'épargne obligatoire, la fraude peut la réduire. Moins d'épargne diminue les capacités de financement de l'activité de production, ce qui se traduit également par une réduction de la croissance des richesses.

C'est donc au nom de critères économiques et de critères de justice sociale que la fraude sociale doit être réduite. Par conséquent, il s'agit de lutter contre ce phénomène pour enrayer sa propagation.

3. La place du dispositif de vérification de l'information dans la lutte contre la fraude sociale

Selon, l'analyse économique des comportements illégaux (Becker, 1968), la fraude est engendrée par des décisions rationnelles : il s'agit de maximiser les gains liés à la fraude sous contrainte de sanctions en cas de contrôle. Allingham et Sandmo (1972), Srinivasan (1972) et Yitzhaki (1974) ont ajouté l'idée selon laquelle il s'agit d'un choix de portefeuille en univers incertain, c'est-à-dire que la fraude ne peut exister que s'il existe un contrôle de la fraude.

Depuis ces travaux, une littérature importante s'est développée concernant les déterminants de la fraude aux prélèvements fiscaux, mais elle se focalise principalement sur le revenu des ménages et les réactions des travailleurs (Alm, 2018). Il existe également des modélisations du comportement des entreprises en matière de fiscalité sur les bénéfices (Marrelli et Martina, 1988 ; Goerke et Runkel, 2006), à propos de la surdéclaration de coût de production pour dépasser les seuils d'imposition (Wang et Conant, 1988), du point de vue du prélèvement à la source de l'impôt sur

le revenu (Yaniv, 1988), sur l'imposition sur les sociétés (Bayer et Cowell, 2009 ; Crocker et Slemrod, 2005 ; Chen et Chu, 2005) et sur la taxe sur la valeur ajoutée (Hashimzade et al., 2010).

En revanche, il existe peu de travaux sur la fraude aux prélèvements contributifs (Bauman et al., 2009). Bien que les deux systèmes se rapprochent sur leur méthode de fonctionnement, il existe néanmoins une différence fondamentale : un prélèvement contributif est une dépense qui donne droit à une prestation individuelle, tandis qu'un impôt sert à financer les dépenses d'intérêt général et dépend des orientations budgétaires de l'État. De plus, la déclaration de l'effectif et la déclaration des cotisations sociales sont des décisions de l'entreprise, qui entrent dans son choix de demander du travail. Malgré ces caractéristiques, il n'existe pas de théorie générale du consentement à la contribution pour les firmes, qui considérerait à la fois la spécificité de la contribution sociale et la demande de travail (Di Porto, 2011).

Cependant, une tentative de modélisation du consentement à la contribution des entreprises à la protection sociale est proposée par Yaniv (1995). Le modèle suggère d'analyser le programme d'optimisation du profit potentiel qui dépend des recettes, de la masse salariale, des contributions patronales et salariales, et du consentement à déclarer les bonnes informations à l'organisme de recouvrement. Ce dernier dispose d'une technique de contrôle imparfaite de la déclaration. La fraude est un comportement rationnel optimal pour l'entreprise. Néanmoins, le modèle se limite à un paramètre incitatif, c'est-à-dire le taux de contribution, pour analyser le consentement.

À notre connaissance, il n'existerait pas de développements récents à ce modèle. Pourtant, des pistes prometteuses permettraient de compléter l'analyse en tenant compte du caractère contractuel de la relation de contribution. À l'instar des modèles principal-agent qui traitent du consentement à payer l'impôt sur le revenu (Reinganum et Wilde, 1985 ; Andreoni et al., 1998), la dimension déclarative peut être considérée comme un aléa moral, c'est-à-dire une information qui est détenue par le déclarant, mais qui n'est pas parfaitement connue par l'organisme de recouvrement. Il s'agit donc de déterminer dans quelle condition un contrat optimal engendre du consentement à la contribution sociale. Mares (2002, 2003) a commencé à poser les questions qui permettraient d'amorcer un modèle plus complet de la relation de contribution : que représente la politique sociale pour les firmes ? Est-ce que l'État social est seulement une contrainte sur les firmes qui vient sous la forme de coûts supplémentaires ou des rigidités du marché du travail non nécessaires ? Ou bien est-ce que la contribution sociale produit des bénéfices immédiats pour les employeurs ? Quand est-ce que les bénéfices offerts par l'État social pour les firmes sont supérieurs aux coûts ?

Les auteurs spécialistes du fonctionnement de la protection sociale ont avancé plusieurs déterminants au consentement au prélèvement social, à la fois du côté des employeurs et du côté de l'organisme de recouvrement, ce qui met bien en lumière la relation entre les deux parties.
Ainsi, les contributeurs ne perçoivent pas le lien entre leurs contributions et les prestations (Auerbach et Kotlikoff, 1987). C'est un problème d'adéquation, parce que les employeurs ne retirent pas un droit direct à prestation, mais ils sont responsables des droits de leurs employés. Les employeurs placent ainsi une faible priorité pour contribuer à la Sécurité sociale relativement aux autres dépenses (Gillion et al., 2000). De plus, le financement de la protection sociale est complexe (McGillivray, 2001), parce qu'il existe plusieurs régimes de Sécurité sociale, plusieurs interlocuteurs de recouvrement, plusieurs taux de contribution en fonction des différents risques couverts par les régimes et plusieurs bases de calcul entre les différentes rémunérations fixes et variables. Par conséquent, plus le système est opaque, plus il est facile pour les employeurs d'utiliser des règles à mauvais escient et de frauder sur la déclaration sociale.
En ce qui concerne l'organisme de recouvrement, il peut tolérer la fraude dans une certaine mesure, soit parce qu'il n'arrive pas à imposer des sanctions efficaces ou bien parce qu'il est incapable de contenir son occurrence (Manchester, 1997). Cette tolérance provient à la fois de son rôle, qui n'est pas celui de contrôler, mais d'assurer une mission de protection, et de ses ressources plus ou moins importantes pour le contrôle.

Même si la littérature théorique ne fournit pas encore de support complet des déterminants du consentement au prélèvement social, les organismes de recouvrement ont développé plusieurs outils organisationnels qui répondent à l'enjeu de réduction de la fraude aux prélèvements sociaux pour assurer la pérennité du système (Enoff et McKinnon, 2010). Cette lutte contre la fraude se fait en partenariat avec les administrations publiques tutelles des organismes de sécurité sociale c'est-à-dire le gouvernement d'un État, qui se matérialise en France par des conventions d'objectifs et de gestion (COG).
La modification des procédures de collecte joue un premier rôle dans la réduction de la fraude. Cela se matérialise dans la simplification des déclarations, par exemple en réduisant le nombre de déclarations sociales à l'administration (Bailey et Turner, 2001). Concrètement, en France, la déclaration sociale nominative (DSN), entrée progressivement en vigueur, remplace, depuis 2017, dix-neuf déclarations sociales aux organismes de sécurité sociale. Cela se réalise également par la simplification des modes de contributions, par exemple en passant à un mode de prélèvement automatique sur les comptes en banque. Dans ce cadre, le nombre d'informations à vérifier se réduit sensiblement, ce qui permet à l'agence de recouvrement de se concentrer sur une unique source d'aléa déclaratif.

La modulation de l'architecture incitative de la protection sociale est une deuxième stratégie pour réduire la fraude. L'architecture des incitations est composée des taux de contribution, des différents types de prestations liées aux contributions, des sanctions encourues en cas de fraude, qu'elles soient pénales, administratives ou financières. En France, les sanctions pénales vont du paiement d'amendes à des peines d'emprisonnement (jusqu'à 3 ans de prison pour les personnes physiques), les sanctions civiles peuvent prendre la forme de redressement forfaitaire, d'annulation des bénéfices de réductions et exonérations de cotisations sociales, et les sanctions administratives sont celles de la dissolution de l'activité et de l'interdiction d'exercer une activité économique voire de l'exclusion des marchés publics et du refus d'aide publique à l'emploi et à la formation professionnelle. Peuvent également s'ajouter des sanctions « complémentaires » par l'affichage public des jugements et l'interdiction d'exercer une fonction publique.

Les stratégies de réduction de la fraude dépendent également de la légitime autorité pour imposer un contrôle efficace (Nyland et al., 2006), c'est-à-dire que l'organisation de sécurité sociale doit être dotée du pouvoir de contrôle et de surveillance des déclarations. C'est dans ce cadre que se situe le dispositif de vérification de l'information. Ce contrôle se caractérise par l'inspection des enregistrements des employeurs, physique ou non, appelée contrôle comptable sur assiette (CCA), l'accès à des informations auxiliaires, comme le compte en banque de l'entreprise, et l'évaluation et la collecte des contributions non payées. Pour faire respecter ce contrôle, les agents doivent se professionnaliser et conduire des contrôles plus opportuns. Cela passe mécaniquement par une allocation plus importante de ressources dédiées au contrôle. Ainsi, les Unions de recouvrement des cotisations de sécurité sociale et d'allocations familiales (Urssaf) s'appuient sur des dispositifs organisationnels, comme la présence depuis 2006 de référents régionaux spécialisés dans la lutte contre la fraude sociale répartis entre un corps de 1 550 inspecteurs du recouvrement et de 220 contrôleurs comptables sur assiette.

La lutte contre la fraude peut aussi se réaliser par une meilleure coordination de l'ensemble des parties prenantes de la protection sociale et de partenaires extérieurs comme les syndicats (Bailey et Turner, 2001), les organisations de santé et les services des impôts (Nyland et al., 2006). En France, la gouvernance de la lutte contre la fraude sociale se situe au niveau administratif, avec la création en 2008 de la Délégation nationale à la lutte contre la fraude (Dnlf), décentralisée au niveau départemental grâce aux Comités opérationnels départementaux antifraude (Codaf). Dans ce cadre, le dispositif de vérification de la déclaration sociale est multiforme et dépend des considérations de plusieurs acteurs.

Enfin, les dispositifs de prévention de la fraude sont à prendre en compte (Goveia, Sosa, 2017). Il s'agit alors d'un exercice de communication pour expliquer et rappeler l'utilité d'un système de protection sociale. Cela peut se traduire par des dispositifs de formation ou de campagne d'informations ciblées (Bailey et Turner, 2001). L'idée étant que la prévention de la fraude doit

permettre d'adopter des attitudes favorables au consentement, c'est-à-dire une moralité envers le prélèvement (Alm et Torgler, 2013), une motivation intrinsèque à payer ses contributions (Feld et Frey, 2002).

4. Concilier la mission de recouvrement avec la mission de vérification de la déclaration sociale

Dans l'ensemble des stratégies et outils à la disposition de l'organisation de recouvrement des contributions sociales destinées à réduire la fraude sociale, nous avons souhaité associer le volet préventif au volet coercitif afin de concilier la mission de recouvrement avec la mission de vérification de la déclaration sociale de l'Acoss.

Dans ce cadre, il s'agit de déterminer la réaction comportementale des cotisants à la suite de la communication de l'architecture du contrôle de la fraude, via une campagne d'information ciblée. Les paramètres du dispositif de vérification de la déclaration sociale ne sont pas modifiés par l'organisation, mais ces paramètres sont rendus visibles aux employeurs. De plus, il s'agit de déterminer si un paramètre particulier du dispositif de vérification de la déclaration sociale influe sur le comportement déclaratif.

Le fait de rendre l'information disponible sur l'architecture du système de vérification de la fraude et les sanctions encourues en cas de contrôle positif doit permettre de provoquer une redéfinition du rapport contractuel du cotisant à la relation de recouvrement. Nous faisons ainsi référence au modèle de Silvers (2012) qui a mis en lumière qu'un contrat optimal d'agence était celui où le principal et l'agent partageaient la même quantité d'informations sur la définition du dispositif de contrôle.

Nous souhaitons caractériser dans quelle mesure une campagne d'information, mobilisant plusieurs types de messages qui contiennent des renseignements sur l'architecture du système de vérification de la déclaration sociale, les modalités de sanctions, et l'utilité de la contribution sociale, permet-elle d'augmenter le consentement au prélèvement social et de réduire la fraude sociale ?

Nous avons expérimenté l'envoi de courriels d'informations ciblées en évaluant son impact sur la déclaration sociale des entreprises. Le protocole expérimental consistait à envoyer des messages aux cotisants sur les aspects qui rentrent dans la décision de déclarer ou non de la masse salariale. La modification de l'ensemble des informations disponibles permet de changer les arbitrages des cotisants et leur décision de déclaration. Chaque groupe faisant l'objet d'un envoi d'informations est constitué de 12 000 établissements du secteur de la construction et 8 000 établissements de l'hôtellerie-restauration, qui sont les deux secteurs dans lesquels la fraude sociale est la plus

répandue, ainsi que 20 000 établissements de l'ensemble des autres secteurs. Les courriels ont été envoyés au cours du dernier trimestre de l'année 2013.

Nous présenterons dans une première partie, les principales pistes de recherche théorique et appliquée permettant d'analyser les effets des stratégies visant à améliorer le consentement à la contribution. En l'absence de travaux portant spécifiquement sur la contribution sociale, notre revue de la littérature tirera les principaux enseignements de la lutte contre l'évasion fiscale. Nous détaillerons ensuite notre méthodologie de recherche ainsi que les principaux indicateurs de mesure. Dans la dernière partie, nous présenterons nos résultats et nous les discuterons.

II. Revue de la littérature

1. Le modèle canonique du consentement au prélèvement obligatoire

Dans cette sous-partie, nous présentons les principales prédictions du modèle canonique du consentement au prélèvement obligatoire, ainsi que ses deux premières extensions qui prennent en compte la spécificité de la contribution sociale des entreprises au financement de la protection sociale des salariés.

a. Les paramètres du modèle canonique

L'étude traditionnelle de la fraude au prélèvement obligatoire s'effectue dans le cadre de l'économie publique (Allingham et Sandmo, 1972 ; Srinivasan, 1973 ; Kolm, 1973 ; Yitzhaki, 1974) en utilisant le modèle de l'économie du crime de Becker (1968) et de l'analyse du choix en univers incertain (Mossin, 1968).

Dans ce modèle, l'individu maximise son utilité espérée tirée d'un choix de déclarer toute ou partie de son revenu imposable à l'administration fiscale, en évaluant les bénéfices d'une fraude réussie, c'est-à-dire une sous-déclaration du revenu non détectée, contre les perspectives risquées de détection et de sanctions. En effet, l'administration ne mesure pas parfaitement les fausses déclarations fiscales, notamment parce qu'elle supporte un coût positif de vérification. La probabilité de détection de la fraude n'est donc pas unitaire. Cependant, lorsque le contrôle permet de détecter une déclaration frauduleuse, le déclarant reçoit notamment une sanction financière.

Pour illustrer cette approche par la théorie de l'utilité espérée, considérons une version simple du modèle canonique. Un individu reçoit un montant fixe de revenu I et il réalise un arbitrage entre ce qu'il compte déclarer ou non aux autorités fiscales. L'individu paye des impôts à un taux t

pour chaque unité R de revenu qui est déclaré, tandis qu'aucune taxe n'est payée sur le revenu non déclaré. Cependant, l'individu peut être contrôlé avec une probabilité fixe p. S'il est contrôlé, alors tout le revenu qui n'est pas déclaré est découvert et l'individu doit payer une amende d'un niveau f pour chaque unité de prélèvement pour laquelle il n'a pas été prélevé.

Si l'individu est contrôlé, le revenu I_c est égal à

$$I_c = I - t.R - f\big(t(I - R)\big)$$

c'est-à-dire le revenu moins les impôts sur le revenu déclaré et retranché du niveau d'amende au titre des impôts non payés.

Si l'individu n'est pas contrôlé, le revenu I_{nc} est égal à

$$I_{nc} = I - t.R$$

c'est-à-dire le revenu moins les impôts sur le revenu déclaré.

L'individu choisit le niveau de déclaration du revenu qui maximise son utilité espérée définie par :

$$EU(I) = p.U(I_c) + (1 - p).U(I_{nc})$$

où $E\big(U(I)\big)$ est l'espérance d'utilité qui ne dépend que du revenu.

L'optimisation donne une condition du premier ordre standard pour une solution intérieure. Étant donné l'hypothèse de concavité de la fonction d'utilité, la condition du second ordre est également satisfaite. Les statiques comparatives sont les suivantes : une augmentation de la probabilité de détection p ou l'augmentation du niveau d'amende f accroît le niveau de déclaration du revenu. L'élasticité du revenu déclaré au taux d'imposition t est plus ambigüe, mais elle tend généralement à être positive (Yitzhaki, 1974).

Par conséquent, cette approche prévoit que l'amélioration du consentement à être prélevé dépend positivement de la capacité de contrôle et de sanctions de l'administration fiscale.

b. L'extension de l'analyse canonique à la fraude aux prélèvements des firmes

Les développements de cette approche se sont principalement concentrés sur les choix opérés par les ménages (Andreoni et al. 1998 ; Alm, 2018 ; Slemrod et Yitzhaki, 2002 ; Marchese, 2004; Franzoni, 2009 ; Sandmo, 2012). Il existe également des modélisations qui ont tenu compte des spécificités des firmes pour évaluer différentes stratégies d'amélioration du consentement au prélèvement (Yaniv, 1995, Goerke, 2014).

La base de ces modèles considère une firme hostile au risque, faisant face à un taux de prélèvement θ imposé sur une assiette (revenu, salaire, stock de capital, profit, etc.) qui se rapporte à un niveau d'activité A (production, emploi, etc.). En tenant compte d'un niveau positif d'activité, le niveau de profit π peut être exprimé par

$$\pi = \pi(A, \theta, \overline{m})$$

où $\overline{m}$ représente un vecteur de paramètres de marché. Lorsque la firme optimise son profit, elle choisit un niveau d'activité A qui donne un profit maximum. En revanche, l'élasticité du profit au taux d'imposition est négative.

La firme est supposée considérer la possibilité de réduire une partie de ses obligations fiscales en sous-déclarant le véritable niveau de son assiette d'imposition, $b(A)$. En notant la déclaration qui dévie de son véritable niveau par S, où $0 \leq S \leq b(A)$, et dans le cas où la firme n'est pas contrôlée le profit sera donc

$$\pi_{nc} = \pi + \theta . S$$

À l'inverse, si la fausse déclaration est détectée, la firme devra payer le montant de la taxe évitée ainsi qu'une amende δ appliquée à ce montant. En cas de contrôle, le profit sera

$$\pi_c = \pi - \delta . \theta . S$$

La fonction d'utilité de la firme U est seulement définie sur les profits et la firme choisit son niveau d'activité ainsi que le niveau de déclaration de telle sorte qu'elle maximise son utilité espérée

$$EU(\pi) = (1 - p) . U(\pi_{nc}) + p . U(\pi_c)$$

où p est la probabilité exogène de détection de la mauvaise déclaration, c'est-à-dire le niveau de vérification de la fraude par l'administration fiscale. La condition du premier ordre pour une solution intérieure implique qu'une fraude au prélèvement existe si $p(1 + \delta) < 1$.

Une augmentation du taux de prélèvement réduit toujours le niveau de déclaration de prélèvement, c'est-à-dire que l'augmentation d'un taux de prélèvement augmente la fraude fiscale des firmes. Ce résultat correspond à la prédiction traditionnelle des modèles canoniques.

De plus, le niveau d'activité est toujours séparable du consentement fiscal, quels que soient les paramètres du contrôle de la fraude, autrement dit le choix optimal du niveau d'activité est le même si la firme déclare toute la base d'imposition ou si elle n'en déclare qu'une partie. Le niveau de consentement au prélèvement est indépendant du niveau d'activité.

L'augmentation de l'amende ainsi que l'augmentation de la probabilité de contrôle réduisent la fraude. Ces paramètres ne modifient pas la décision du niveau optimal d'activité (Goerke, 2014).

c. La prise en compte de la spécificité de la contribution sociale des entreprises

Les développements théoriques du précédent modèle ont principalement concerné la fraude fiscale des firmes. À notre connaissance, Yaniv (1995) serait la seule proposition de modélisation de la fraude aux prélèvements sur la masse salariale.

Dans son modèle, il considère un employeur qui fait face à un salaire fixe par travailleur w sur une période de temps donnée et qui doit contribuer à la protection sociale par une proportion θ_p de sa masse salariale et de retenir également à la source une proportion θ_s de sa masse salariale pour le compte de ses employés. Il n'y a pas de plafond de contribution.

Le profit de l'employeur en cas de consentement total à la contribution sera

$$\pi(L, \theta_p, w) = V(L) - (1 + \theta_p).w.L$$

où L représente l'effectif et $V(L)$ les recettes totales. Cependant, si l'employeur sous-déclare sa masse salariale d'un montant $S \leq w.L$, ses profits nets seront donnés en cas de non-détection par l'équation $\pi_{nc} = \pi + \theta.S$. Et en cas de détection par l'équation $\pi_c = \pi - \delta.\theta.S$ où $\theta = \theta_p + \theta_s$.

En appliquant le même cadre de la maximisation de l'utilité espérée, la principale prédiction de ce modèle est qu'une augmentation des contributions patronales ou salariales décourage la sous-déclaration.

2. Les enrichissements du modèle

Les principales prédictions de cette approche par l'utilité espérée ne sont néanmoins pas confirmées par les preuves empiriques (Alm, 2018). En effet, empiriquement, le consentement tend à diminuer avec de plus hauts taux de prélèvement.

De plus, les individus ne réalisent pas uniquement leurs choix à partir de considérations financières. Le modèle prédit ainsi que les individus devraient rationnellement sous-déclarer leur revenu. Or, dans les faits, le phénomène de sous-déclaration n'est pas aussi répandu que la théorie l'envisage et la majorité des individus payent l'intégralité de leurs impôts, quel que soit le niveau d'incitations auquel ils font face.

Soulignons également que le faible nombre de travaux théoriques sur le comportement des firmes en matière de contribution sociale ne permet pas de produire de prédictions plus précises quant aux stratégies à mettre en œuvre pour améliorer le consentement des employeurs.

Enfin, la contribution sociale peut s'apprécier comme une relation contractuelle, qui engendre des engagements de réciprocité avec l'organisme de collecte des contributions sociales et l'employeur.

L'ensemble de ses limites plaide pour un cadre théorique approfondi pour analyser les déterminants théoriques de la réduction de la fraude aux prélèvements sociaux.

a. Une première voie d'enrichissement : l'approche contractualiste du prélèvement

Une extension de l'analyse théorique réside dans le cadre de l'utilité espérée et consiste à étudier le contrôle endogène de la fraude au prélèvement. En effet, dans l'environnement du prélèvement, il existe un second acteur pour lequel le comportement n'est pas modélisé : il s'agit de l'administration fiscale.

Théoriquement, il est possible de considérer que le prélèvement est une relation contractuelle et que les deux pôles de la relation agissent comme dans une relation d'agence où se situe un aléa moral sur le niveau de déclaration des ressources de l'agent. Nous considérons ainsi que le développement proposé par Reinganum et Wilde (1985) semble pertinent pour analyser la relation de prélèvement. Ce modèle est toutefois destiné à l'analyse de l'imposition sur le revenu des ménages. Nous décrivons dans un premier temps le modèle et ses conclusions et nous expliquons en quoi il peut être judicieux de recourir à ce modèle pour l'analyse du recouvrement de la contribution sociale et de l'amélioration du consentement au prélèvement social.

Le point de départ de leur analyse est que la déclaration d'un montant imposable est par nature volontaire. Par ce caractère, ils étendent l'analyse traditionnelle, dans laquelle le consentement au prélèvement est une décision en univers incertain. En effet, dans le modèle canonique, les individus font face à une probabilité de détection indépendante du montant déclaré, c'est-à-dire que c'est une probabilité non conditionnelle à un seuil de détection.

Dans ce modèle principal-agent appliqué au consentement à prélever, l'agent désigne l'individu qui est taxé et le principal désigne l'individu qui prélève la taxe. Dans cette relation d'agence, l'aléa moral réside dans le montant imposable déclaré par l'agent au principal.

L'agent « observe » sans coût un certain montant w sur lequel est appliqué un prélèvement et il fournit une déclaration d'un certain montant x. Le principal observe le montant à prélever w s'il effectue un contrôle $p(x)$ à un certain coût c. S'il n'y a pas de contrôle, le transfert de l'agent vers le principal correspond au niveau de prélèvement $t(x)$. En revanche s'il y a un contrôle, il peut y avoir une amende si le montant déclaré n'est pas le bon en fonction de $f(x, w)$.

L'agent a un niveau d'utilité espérée, qui dépend du niveau du montant déclaré ainsi que du prélèvement :

$$u_a(x, w) = [1 - p(x)].[w - t(x)] + p(x).[w - f(x, w)]$$

Il maximise son utilité sous contrainte du prélèvement, de l'amende en cas de fraude et de la probabilité de contrôle.

Le principal a un niveau d'utilité espérée, qui dépend du transfert ainsi que du coût du contrôle et du niveau d'amende lorsqu'il y a une fraude :

$$u_p(x, w) = [1 - p(x)].t(x) + p(x).[f(x, w) - c]$$

En tenant compte de la déclaration optimale de l'agent, le principal sélectionne la fonction d'imposition, la fonction d'amende et la fonction de contrôle. Il détermine enfin quel est le seuil de détection de la fraude, c'est-à-dire le montant en dessous duquel, le montant déclaré est considéré comme frauduleux.

Le modèle prédit qu'une technologie de contrôle aléatoire et sans amende permet de consentir au prélèvement. En revanche, un seuil d'audit trop bas et sans amende est inefficace. Enfin un contrôle total sur tous les niveaux de revenus est jugé efficace.

À notre connaissance, il n'existe pas de modèle principal-agent de la relation de recouvrement. Cependant, nous pouvons aisément adapter ce cadre en considérant les éléments suivants.

D'une part, nous pouvons considérer que l'Acoss, l'individu qui collecte une certaine quantité de ressources, dont l'objectif est de financer les prestations sociales est le principal.

D'autre part, nous pouvons considérer que l'individu qui doit contribuer au financement de la protection sociale est l'agent. Dans le cadre du prélèvement social, ce sont les producteurs de biens et services qui contribuent. Il est possible de scinder en trois dimensions le comportement d'un producteur : le comportement d'offre sur le marché des biens et services ; le comportement de demande des facteurs de production, notamment le facteur travail ; le comportement de déclaration sur la base du volontariat de la masse salariale, qui constitue l'assiette du prélèvement. L'objectif du producteur est de maximiser son profit sous une contrainte de coût, laquelle est multiple : coût du travail en tenant compte des salaires versés aux employés ; coût du prélèvement social sur la masse salariale ; coût de la sanction s'il y a fraude.

En tenant compte de cette dernière caractéristique de la fonction de coût, on peut affiner le modèle en précisant une « séquence comportementale » qui inclurait la caractéristique de fraude à la fois dans la demande de travail et dans la déclaration des revenus. Ainsi, un producteur peut déclarer correctement sa masse salariale s'il demande du travail formel et qu'il déclare correctement l'assiette de contribution sociale. À l'opposé, il peut choisir de déclarer en réalisant des erreurs s'il demande du travail informel et sous-déclare la contribution. Dans ce cas, nous désignons la situation de fraude sociale, qui est la rencontre conjointe entre le fait d'embaucher un individu en dehors de la législation qui encadre les conditions de travail et le fait de ne pas déclarer au collecteur la réalité de l'activité. Enfin, un producteur peut également déclarer avec erreur lorsqu'il demande du travail formel. Ce dernier cas peut traduire une situation de mauvaise déclaration, sans que cela soit une fraude consciente et choisie, ou à l'inverse être un moyen de ne pas consentir au prélèvement.

Pour le producteur, c'est la décision de déclarer correctement ses informations sociales qui permet d'agir sur le marché du travail et sur le marché des biens. Il existe donc un aléa moral sur le bon niveau de déclaration sociale du producteur, c'est-à-dire une incertitude sur le fait de bien déclarer ses informations sociales.
C'est pourquoi le collecteur de la cotisation sociale possède un dispositif de vérification de la déclaration sociale. Ce dispositif est coûteux pour plusieurs raisons : le collecteur doit se déplacer physiquement pour réaliser une vérification de l'activité ; le contrôle peut être long s'il s'agit de visiter les lieux de production répartis à des endroits différents ; il faut rémunérer les contrôleurs. Il s'agit donc d'un contrôle imparfait de l'ensemble des producteurs. Lorsque le contrôle détecte un producteur en situation de fraude, cela déclenche une sanction financière, le plus souvent, la régularisation des sommes dues, à laquelle s'ajoute une amende pour fraude et le paiement d'indemnités pour les employés en situation de travail informel. Il peut également s'agir de sanctions non financières qui entraînent la cessation d'activité et jusqu'à l'emprisonnement du fraudeur. L'existence de ces sanctions joue le rôle d'incitations à modifier le comportement des producteurs, de façon ex post, après avoir été détectés comme fraudeurs, mais également de façon ex ante, avant le déclenchement d'un contrôle.

La séquence de choix est la suivante : l'agent (le producteur) va choisir son comportement de déclaration qui maximise son profit en tenant compte du risque qu'il prend à ne pas faire une déclaration correcte ; le principal (le collecteur) va choisir son dispositif de vérification du prélèvement et le taux de prélèvement qui permet de financer ses dépenses sociales. Le meilleur contrat est celui qui permet d'aligner les intérêts des deux pôles de la relation.
Dans ce contexte, la fraude peut exister si les incitations ex ante n'ont pas un impact suffisamment important sur la modification du comportement de déclaration sociale, ou bien si la technologie de vérification de la déclaration sociale n'est pas assez précise.
Dans le cadre de ce chapitre, nous ne prenons pas le parti de réaliser un modèle théorique enrichi. Ces pistes de réflexion permettraient cependant de contribuer à une analyse plus approfondie de la relation de prélèvement social.

Bien qu'enrichi, un tel modèle contient également les principales limites du modèle canonique, à savoir que les individus réagissent essentiellement aux incitations financières en payant leur contribution seulement parce qu'ils craignent la détection et la sanction. De plus, la prédiction d'une élasticité positive de la déclaration à une hausse du taux de prélèvement est de nouveau un résultat important de ce modèle, ce qui continue d'être en inadéquation avec les preuves empiriques (Alm, 2018).

b. Second enrichissement du modèle : économie comportementaliste

Un second type d'enrichissement du modèle de fraude aux prélèvements prend en compte les développements de l'économie comportementale (Pickhardt et Prinz, 2014), définie comme une approche qui utilise les méthodes et les démonstrations des autres sciences sociales, en particulier la psychologie, pour analyser différemment les prises de décisions individuelles et collectives par rapport au cadre standard (Camerer et al., 2004)[5].

Le courant comportementaliste a adopté des déviations aux hypothèses standards pour analyser des processus de décision individuelle plus proche de ce qui est observé empiriquement. Néanmoins, lorsque le courant se concentre sur le point de vue individuel, les prédictions des modèles ne changent pas véritablement du cadre standard, notamment parce que les gains associés à la fraude aux prélèvements dépendent toujours du taux de prélèvement (Alm, 2018).

En ce qui concerne l'inclusion de l'environnement social dans la décision de déclarer, c'est la notion de « moralité fiscale » qui s'est développée. En effet, bien que les modèles canoniques et comportementalistes prédisent une large diffusion de la fraude, en réalité les individus décident de volontairement payer leurs impôts et leurs contributions sociales, quel que soit le niveau de prélèvement. Slemrod (2007) formulait en ces termes que « l'énigme de la fraude n'est pas de savoir pourquoi les gens fraudent, mais bien pourquoi les gens ne fraudent pas ».
Une motivation intrinsèque à payer les prélèvements existerait (Frey, 1994, 1997), c'est-à-dire qu'il existe une tendance naturelle, une inclination spécifique aux individus pour contribuer à un bien commun. Cette motivation intrinsèque serait ainsi façonnée par des valeurs culturelles, des normes sociales, des institutions sociales. La motivation intrinsèque ou moralité fiscale serait ainsi un important déterminant du consentement au prélèvement (Torgler, 2008). Par ailleurs, prendre cette dimension « sociale » et les relations qu'entretiennent les individus avec le collecteur du prélèvement a permis d'aboutir à des prédictions théoriques plus en adéquation avec les preuves empiriques et a notamment contribué à développer l'idée que l'amélioration globale de la qualité

[5] Dans le cadre standard, l'individu est rationnel, possède un désir sans limites, et est purement conduit par son intérêt personnel. Dans le cadre comportementaliste, trois déviations sont principalement utilisées pour décrire l'individu (Congdon et al., 2011) : l'individu ne réalise pas de parfaite optimisation parce que sa rationalité est limitée par de mauvaises perceptions des coûts et des bénéfices réels de ses actions et qu'il ne peut pas réaliser tous les calculs nécessaires pour prendre ses décisions ; l'individu n'a pas une maîtrise totale de lui-même et cela le conduit à faire des choix incohérents dans le temps, c'est-à-dire que ses choix sont biaisés par la formulation des informations dont il dispose ; ses préférences sont influencées par des considérations non financières, notamment celles qui sont liées à l'environnement social, comme l'équité (l'idée d'impartialité, ce qui est juste), l'altruisme (la disposition à manifester de l'intérêt, de la générosité envers autrui), la réciprocité (l'idée qui implique un échange d'actes ou de sentiments entre deux individus), l'empathie (la capacité de se mettre à la place de l'autre, de ressentir la même chose que lui, de s'identifier à lui), la confiance (sentiment de sécurité, d'assurance, d'espérance convaincue inspiré par quelqu'un).

de la relation entre contribuable et collecteur de taxes permet d'améliorer le consentement et de réduire la fraude (Alm, 2018).

À notre connaissance, hormis Alm et McClelland (2012), il n'existe pas d'autres travaux qui porteraient sur la moralité fiscale des entreprises. Ces auteurs l'expliquent notamment par le fait qu'il n'existerait pas de mesure de la moralité fiscale au niveau des firmes ce qui empêcherait de modéliser des déterminants spécifiques de leur moralité fiscale.

Leur contribution est plus économétrique que théorique, mais leur méthodologie permet d'appréhender le phénomène de moralité fiscale pour les firmes. Ils ont couplé deux enquêtes, la Business Environnement and Enterprise Performance Survey (BEEPS) et la World Enterprise Survey (WES) et ont cylindrées les réponses aux questions sur 8500 entreprises réparties sur 34 pays (principalement Europe centrale et de l'Est, Corée du Nord, Irlande, Espagne et Allemagne) sur une période de 8 ans.

Le consentement au prélèvement, c'est-à-dire la déclaration à l'administration, est mesuré par le taux de vente qu'une firme typique doit déclarer dans son pays, dépend des actions coercitives de l'administration fiscale, c'est-à-dire si la firme a reçu ou pas une visite de l'administration dans les 12 derniers mois, des obstacles fiscaux, c'est-à-dire les difficultés à payer les taxes, et de sa structure de propriété (nationale, étrangère, publique), de son secteur d'activité, de ses ventes, du sexe du propriétaire, du pays et de l'année de l'enquête.

À cause d'un potentiel biais d'endogénéité de la variable de coercition administrative, les auteurs ont modélisé cette variable à partir des coupures de courant et d'alimentation en eau, l'idée étant qu'un réseau énergétique robuste est un indicateur valable de l'efficacité du contrôle des administrations fiscales, tandis que cette robustesse ne doit pas avoir d'effet sur le niveau de déclaration des ventes de la firme.

Les obstacles pour payer les taxes dépendent de plusieurs facteurs. Par exemple, l'incidence d'un prélèvement sera plus contraignante pour une petite firme avec peu de ressources qu'une grande firme avec plus de ressources dédiées au paiement des charges. En revanche, plus une firme gagnera en taille, plus les prélèvements auxquels elle fera face seront diversifiés.

De plus, la moralité fiscale est un déterminant supplémentaire de l'obstacle perçu à payer les impôts. Selon les auteurs, la moralité fiscale est l'attitude envers (1) le prélèvement en général et envers (2) le fait de payer des taxes. Une firme avec une faible hostilité envers le paiement des taxes peut considérer les prélèvements dans leur globalité négativement, et réciproquement. Une haute moralité fiscale pour une firme sera de considérer les taxes et le paiement des taxes comme de faibles obstacles à la conduite de l'activité. Comme la moralité fiscale n'est pas directement mesurable dans le questionnaire, les auteurs utilisent deux questions indirectes qui renvoient aux deux déterminants de la moralité :» le taux de prélèvement est-il un obstacle » (i.e. le fait de payer

des taxes) et « l'administration fiscale est-elle un obstacle » (i.e. l'ensemble des taxes). Les autres variables observables sont utilisées pour modéliser l'obstacle à payer des taxes.

Le résultat de leur étude montre que plus la moralité fiscale est élevée, plus le taux de déclaration des ventes est élevé, ce qui montre un lien positif entre le fait de considérer que l'évasion fiscale est mauvaise et le fait de moins frauder.
Leurs estimations permettent également de montrer que d'autres stratégies alternatives à la coercition sont efficaces pour améliorer le niveau de déclaration, notamment celles qui augmentent la moralité fiscale : mettre l'accent sur le fait que le paiement des impôts est la bonne chose à faire, publier le nom des fraudeurs, améliorer la liaison entre prélèvement et service public, encourager la participation des usagers aux arbitrages budgétaires, s'occuper des iniquités de traitement fiscal et éviter que la fraude soit tolérée.
Leurs conclusions indiquent que la moralité fiscale s'applique tant pour les ménages que pour les firmes. Si les contrôles fiscaux étaient considérés comme le « bâton » de la coercition, alors ces stratégies alternatives devraient être considérées comme la « carotte » pour améliorer le consentement au prélèvement.

Sans aucun doute, il est possible d'appliquer ce concept de moralité au prélèvement à la contribution sociale. Comme nous l'avons précédemment évoqué, la conception de la protection sociale est que les individus contribuent pour ouvrir des droits sociaux pour eux-mêmes, mais également pour les autres, selon un principe de solidarité, c'est-à-dire une relation entre des personnes unies par un sentiment de communauté d'intérêts qui les pousse à se porter aide mutuelle. On comprend aisément la dimension collective de la protection sociale. Dans ce contexte, la moralité doit également jouer un rôle dans le consentement à la contribution sociale. Adopter des stratégies alternatives à la modulation des incitations pour permettre de réduire des comportements illégaux est une sérieuse voie de recherche à prendre en considération.

c. D'autres extensions possibles pour aboutir à un cadre unifié du consentement au prélèvement social

Nous souhaiterions aborder dans cette section, des spécificités du comportement de demande de travail afin d'aboutir à un cadre théorique plus général sur les stratégies à adopter pour réduire la fraude aux prélèvements sociaux.
En effet, comme nous l'avons établi en introduction et évoqué dans la modélisation canonique enrichie par l'approche contractualiste, la fraude sociale a deux versants : la rémunération et l'effectif. La non-déclaration de l'effectif peut être définie comme toute activité rémunérée qui est légale au regard de leur nature, mais qui n'est pas déclarée à l'administration. Cette définition

rentre ainsi dans le champ du travail dissimulé, ou travail informel (Hart, 1972). C'est un processus qui engage les employeurs et les employés volontairement à cause du gain potentiel pour échapper aux taxes et aux contributions sociales.

Cependant, les nombreux développements théoriques pour analyser le secteur informel et les moyens de restreindre son étendue ont essentiellement porté sur l'offre de travail (Sandmo, 2005). Les quelques contributions comme celle de Fugazza et Jacques (2003) considèrent que les firmes peuvent être dans l'une ou l'autre face de l'économie sans être à la fois dans les deux, c'est-à-dire réaliser des activités légales tout en employant de la main-d'œuvre illégalement. Plus rares encore sont les contributions qui portent leur attention sur la spécificité de la contribution sociale.

Néanmoins, nous avons recensé deux documents de travail qui posent les bases d'un approfondissement du modèle de fraude au prélèvement qui tiendrait compte d'éléments de l'économie industrielle afin de modéliser le comportement de demande de travail et d'évaluer certaines prédictions au regard de l'amélioration du consentement au prélèvement en général des firmes et de leurs contributions sociales en particulier.

Vlassis et Mamakis (2014) considèrent un duopole à la Cournot où deux firmes se font concurrence sur les quantités. L'une d'entre elles choisit de déclarer sa masse salariale, contribue à la Sécurité sociale en payant une cotisation sur sa masse salariale, k, tandis que l'autre ne déclare pas sa masse salariale. Les deux firmes payent un impôt sur leurs bénéfices d'un niveau t. Dans leurs hypothèses, la firme qui déclare sa masse salariale bénéficie d'une déduction fiscale. De plus, le salaire est exogène, ainsi que les décisions de déclaration. Dans ce modèle, il n'y a pas de précision au sujet du contrôle de la masse salariale.

On peut écrire les profits des entreprises de la façon suivante :

$$\pi_1 = (p(Q) - (1-k).w).q_1 - t.[(p(Q) - (1-k).w).q_1]$$

$$\pi_2 = (p(Q) - w).q_2 - t.[p(Q) - w.q_2]$$

En optimisant les profits et en tenant compte des solutions du premier ordre, il est possible de donner la condition de répartition des parts de marché entre les firmes. La condition de répartition égalitaire du marché est donnée par :

$$t = \frac{k}{1+k}$$

Ainsi si $t > \frac{k}{1+k}$, alors $q_1 > q_2$, ce qui implique que la firme qui déclare sa masse salariale et paye ses contributions sociales possède plus de part de marché. À l'inverse, si $t < \frac{k}{1+k}$, alors $q_1 < q_2$, ce qui implique c'est la firme qui ne déclare pas sa masse salariale qui possède plus de part de marché.

La conclusion de ce modèle est qu'une augmentation de la fiscalité sur les bénéfices peut engendrer une diminution du travail dissimulé.
De nouveau, seule l'incitation financière modifie la décision de déclarer sa masse salariale. Néanmoins, cette première contribution permet de proposer un cadre renouvelé de l'analyse de la fraude au prélèvement des firmes.

Le travail réalisé par Diaye et Koshevoy (2006) est celui qui a développé le plus loin l'analyse de l'interaction entre la fraude à la dissimulation de l'activité et la demande de travail. Leur modélisation croise l'économie industrielle et l'économie des contrats en proposant un modèle de concurrence oligopolistique à la Cournot croisé avec un modèle principal-multiagents afin de prédire les effets d'une augmentation de la détection de l'emploi de travailleurs illégaux sur le choix de demander du travail.
Dans ce modèle, on considère une économie de marché avec un secteur légal et un secteur illégal, où tous les marchés sont à l'équilibre sauf celui du travail. On considère un oligopole de firmes. Les firmes du secteur légal peuvent employer des travailleurs du secteur légal en quantité L_l à un salaire w et des travailleurs du secteur illégal en quantité L_{il} à un salaire b. Il existe un degré de substituabilité entre les types de main-d'œuvre qui engendre un coût de transaction, c'est-à-dire que lorsque les firmes souhaitent employer un travailleur illégal à la place d'un travailleur légal, elles doivent verser une compensation pour améliorer les compétences de ce dernier.
Il existe une variable aléatoire A qui vaut 1 si la firme choisit d'employer des travailleurs illégaux et 0 sinon. Il existe également une variable aléatoire X qui vaut 1 si la firme est détectée lorsqu'elle emploie des travailleurs illégaux, 0 sinon. On note $\mu = p(X = 1)$ la probabilité de détecter l'emploi de travailleurs illégaux. Si la firme est détectée, elle subit une amende t dont le montant dépend de la quantité utilisée de main-d'œuvre illégale. Dans ce modèle, il existe également un coût social pour la firme s qui inclut les conséquences financières, par exemple si les consommateurs boycottent les produits de la firme.
La firme optimise le profit suivant selon les quantités de travail légales ou illégales :

$$\pi = p(Q).f(L_l, L_{il}) - w.L_l - b.L_{il} - \mu(t + s)$$

Le salaire d'équilibre des travailleurs illégaux est égal au salaire légal moins le coût anticipé de la détection de la fraude, c'est-à-dire que la firme transfère le coût du contrôle $\mu(t + s)$ à leurs travailleurs illégaux en réduisant leur salaire net illégal.
Cette proposition étant déterminée, les auteurs analysent ensuite l'état du marché en oligopole et utilisent le cadre d'un modèle principal-multiagent sur la relation de détection de l'activité illégale.
Les agents sont les firmes, qui par hypothèse ne coopèrent pas. On note y_l la quantité d'output réalisée par la main-d'œuvre légale et z_{il} la quantité d'output réalisée par la main-d'œuvre illégale.

Les coûts associés à la production prennent en compte le coût du travail et les autres taxes, et le coût du travail légal est supérieur au coût du travail illégal.

Le principal est l'administration. Elle ne connaît pas le niveau d'output légal et illégal de chaque firme ni leurs coûts de production. En revanche, elle observe l'équilibre sur le marché ainsi que la demande inverse. Le principal peut détecter le niveau d'output illégal en comparant le prix de marché associé à une production légale déclarée, au prix optimal de marché associé à une production légale et illégale.

Il existe plusieurs mécanismes de détection, notés $D = \{d, \mu\}$, où d est le coût de vérification et de détection de la production illégale et μ la probabilité qu'une firme soit détectée. Dans ce cadre, elle n'est plus constante mais elle dépend des quantités d'output, ainsi $\mu = \mu(\sum(y_l + z_{il}))$. Le choix d'un dispositif du contrôle est dans ce modèle exogène.

Finalement, la firme choisit la quantité d'output légale et illégale qu'elle déclare. Elle choisit également la probabilité à laquelle elle « veut » être détectée. En notant $F = \rho \sum z_{il}$ le seuil de détection où ρ est la partie de la production déclarée qui est contrôlée par le principal, on peut écrire que la probabilité de détection est :

$$\mu\left(\sum(y_l + z_{il})\right) = \begin{cases} 0 \; si \; F < F_1 \\ \dfrac{F - F_1}{F_2 - F_1} \; si \; F_1 \leq F \leq F_2 \\ 1 \; si \; F > F_2 \end{cases}$$

Ce qui implique que la probabilité de détection est nulle si la fraude est inférieure à un certain seuil et totale si la fraude est supérieure à un second seuil de détection. Entre ces deux seuils, le contrôle est imparfait à cause des stratégies de dissimulation.

Le profit de chaque firme est le suivant

$$\pi = p(y_l + z_{il}).(y_l + z_{il}) - c_1.y_l - c_2.z_{il} - \frac{F - F_1}{F_2 - F_1}.z_{il}.(t + s)$$

Les auteurs concluent avec trois propositions. Premièrement, plus le seuil d'une détection totale de la production illégale F_2 est bas plus la quantité d'output illégale z^* est basse. En effet, afin d'éviter de se faire détecter la firme réagit en diminuant la quantité de production illégale. Deuxièmement, la probabilité optimale de détection μ^* dépend négativement de la part de l'activité contrôlée par l'administration ρ et négativement du seuil de détection totale de la production illégale F_2. En d'autre terme, du point de vue de la politique publique, si la probabilité de détection est endogène, vouloir un objectif de contrôle de l'activité illégale en augmentant le nombre de contrôles ou en augmentant la sanction financière sera un échec. En effet, les firmes réagiront à cette politique en continuant d'utiliser le marché illégal tant qu'il demeure profitable.

En conclusion, dans ce modèle, le contrôle ne permet pas d'améliorer le consentement au travail non dissimulé.

Cette revue de la littérature théorique permet de prendre la mesure des différents aspects de la fraude aux prélèvements sociaux et des différentes stratégies à adopter pour paramétrer les dispositifs concourant à la réduire. D'une part, le dispositif de vérification de l'information peut jouer un rôle dans l'amélioration du consentement. D'autre part, d'autres éléments sociaux sont à prendre en compte afin de rendre plus complète la perception du consentement.
Nous pouvons conclure cette partie théorique en proposant qu'il n'existe pas encore de cadre théorique complet qui permettrait d'analyser les déterminants qui réduiraient la fraude au prélèvement social, mais que des pistes prometteuses offertes aux théoriciens sont ouvertes pour y parvenir.

3. Les méthodes appliquées pour améliorer le consentement au prélèvement

Dans cette partie, nous souhaitons aborder les apports des travaux empiriques conduits pour mesurer les effets de dispositifs visant à améliorer le consentement au prélèvement. Nous décrivons les principales méthodologies utilisées pour mener ses travaux et nous présentons les principaux résultats des recherches menées sur le consentement au prélèvement des firmes.

a. L'évolution des méthodes d'évaluation du consentement au prélèvement

Le fait de payer ses prélèvements est le résultat d'un consentement quasi volontaire et de la coercition. Néanmoins, il est évident que les niveaux élevés de consentement au prélèvement ne sont pas seulement dus à la coercition. Les recherches empiriques ont testé la validité des déterminants du consentement dans un cadre réel. Le problème inhérent à la recherche empirique est la difficulté d'obtenir des réponses honnêtes à propos d'un comportement malhonnête.
C'est pour cette raison que dans un premier temps, les recherches par questionnaires se sont focalisées sur les attitudes plutôt que sur le comportement réel (Slemrod, 2007). Par exemple dans le World Value Survey : « dites-moi pour ces actions suivantes si vous pensez qu'elles peuvent être toujours justifiées, jamais justifiée ou quelque part entre les deux : recourir à la fraude fiscale si vous en avez l'occasion ». Cette littérature permet d'apprécier les déterminants du consentement, mais elle souffre d'une limite majeure : les attitudes que les individus relèvent dans leurs réponses ne reflètent pas forcément leur comportement réel (Ajzen et Fishbein, 1977).

Par conséquent, les chercheurs ont employé les méthodes expérimentales pour dépasser cette limite et identifier des effets causaux des déterminants du consentement. Dans les premières expérimentations fiscales, les liens de causalité ont été identifiés en laboratoire. L'artificialité du

laboratoire a permis de contourner le problème d'avoir des réponses honnêtes à propos d'un comportement illégal en travaillant dans un cadre hypothétique. Ces expérimentations ont été classées en deux champs de recherche (Torgler, 2002) : « celles qui parlent aux théoriciens » afin de tester les prédictions des modèles théoriques sur les principaux déterminants du consentement (taux de prélèvement, sanctions, probabilité de détection de la fraude) ; « celles qui sont à la recherche des faits » afin de tester d'autres déterminants (le rôle des biens publics, le rôle de la morale et des facteurs sociaux) en se focalisant principalement sur le revenu des ménages.
Néanmoins, ces deux types d'expérimentations contiennent un certain nombre de limites. Les étudiants répondent à des incitations, mais ils ne sont pas de vrais contribuables et ils peuvent ne pas avoir eu d'expérience avec les conditions réelles du prélèvement. De plus, les échantillons sont généralement réduits ce qui rend le pouvoir statistique des estimations faible et les résultats peu fiables. Les conditions d'artificialité peuvent également être contre-productives dans la mesure où le comportement en laboratoire peut être différent à ce qui est observé sur le terrain dans un contexte social réel et lorsqu'il faut faire face aux véritables conséquences d'un comportement illégal. Par conséquent, ces limites ont conduit les expérimentations fiscales en laboratoire à avoir peu d'impact sur les décisions politiques (Mascagni, 2018).

Des expérimentations de terrain ont donc été conduites afin d'atteindre un plus fort gain de validité externe. Les premières ont répliqué les méthodes employées en laboratoire, avec les mêmes limites (Cummings et al., 2009). La seconde génération des expérimentations de terrain peut être labélisée comme « chuchotant à l'oreille du prince » (Kagel et Roth, 1995) parce que les expérimentateurs souhaitent modifier la politique publique.
Les participants à ces expérimentations sont de vrais contribuables, les données proviennent de l'administration, et l'expérimentation est réalisée dans un environnement réel. Le traitement consiste à envoyer des informations destinées à modifier les déterminants du consentement comme la perception d'une probabilité de contrôle, l'équité et la moralité fiscale. Ces informations sont traditionnellement envoyées par courrier dont le contenu met en lumière les contrôles conduits par l'administration fiscale, la sévérité des amendes que doivent supporter les fraudeurs décelés, ou encore la primordialité des impôts dans le financement des services publics.
Ces expériences sur le terrain, en étudiant les facteurs comportementaux du contribuable, peuvent potentiellement fournir des conseils sur la manière d'augmenter les recettes fiscales de façon rentable en changeant les perceptions et les informations disponibles aux contribuables. Ceci est particulièrement important, car augmenter de manière crédible la coercition est coûteux et pas toujours souhaitable. Avec cette information, les décideurs peuvent évaluer leurs options politiques et paramétrer leurs stratégies de contrôle par rapport aux déterminants qui génèrent le plus de réponses favorables à un coût raisonnable. Enfin, les données récoltées durant une

expérimentation de terrain peuvent être utilisées pour estimer l'ampleur de la fraude aux prélèvements.

Cependant, le développement de ces expérimentations de terrain dans le champ des prélèvements obligatoires est relativement récent, à cause notamment du coût logistique élevé que ces expérimentations impliquent (Mascagni, 2018). De plus, elles sont souvent instantanées, c'est-à-dire qu'elles n'interviennent qu'à un certain moment sans qu'il soit aisé de déterminer les effets durables de ces dispositifs (Alm, 2018).

b. Les principaux résultats des expérimentations analysant le comportement de déclaration des firmes

Les expérimentations de terrain ont principalement porté sur le comportement de déclarations de revenu par les ménages (pour une revue de la littérature, se référer à Alm, 2018 et Mascagni, 2018). Plus rares sont les études qui portent sur le comportement des firmes.

Nous nous restreignons à réaliser une revue des études employant une méthodologie de recherche similaire à celle employée dans ce chapitre, c'est-à-dire des expérimentations de terrain qui ont testé les prédictions du modèle canonique, enrichi du consentement au prélèvement via l'envoi d'informations à propos de la fiscalité des firmes. Cette restriction est effectuée pour plusieurs raisons : cela permet de retirer du champ de la revue d'une part, les études dont l'objet d'étude portait sur le comportement de consentement au prélèvement des ménages ; et d'autre part, celles qui utilisent des méthodes d'expérimentations naturelles ou non expérimentales. Le lecteur souhaitant apprécier les résultats de ces recherches peut se rapporter aux revues plus complètes (Mascagni, 2018 ; Torgler, 2003).

À notre connaissance, la première investigation expérimentale de l'envoi d'information à des entreprises a été réalisée par Barak (2011). Dans cette expérimentation conduite sur des firmes israéliennes, deux types de messages ont été envoyés par tirage au sort afin d'améliorer leur consentement à l'imposition sur les bénéfices : un message de dissuasion de la fraude exposant les nouveaux dispositifs de vérification des comptes et les modalités de sanction des entreprises effectuant de mauvaises déclarations de leurs bénéfices ; un message selon lequel « tout le monde souffre de la fraude », afin de sous-entendre que le prélèvement correspond à la norme sociale. 731 firmes ont reçu le premier message et 732 autres ont reçu le second, tandis que 2 932 firmes constituaient le groupe témoin. Dans ce protocole, le groupe témoin n'a pas reçu de courrier. Les résultats de cette première expérimentation sont relativement négatifs : le chiffre d'affaires déclaré a augmenté tendanciellement pour les deux groupes traités, mais en moyenne celui du groupe traité n'est pas significativement différent de celui du groupe témoin ; le niveau de prélèvement de la taxe sur la valeur ajoutée (TVA) n'est également pas impacté par l'expérimentation ; enfin, le

nombre de demandes pour des déductions de la TVA a augmenté pour les deux groupes, et le groupe qui a reçu le message de persuasion morale a significativement demandé davantage de déductions fiscales. L'auteur conclut que ce type de dispositif peut engendrer des comportements contre-productifs.

Au Chili, Pomeranz (2015) a réalisé une expérimentation sur 102 000 entreprises de tous secteurs confondu. Elle portait sur l'envoi d'un message d'information sur les règles en matière de déclaration de TVA, qui contenait également une description des règles en matière de vérification des comptes. Ce message a eu un effet positif sur le montant des prélèvements pour le compte de la TVA, et cet effet est plus important pour les petites entreprises.

En 2011, dans la région de Styrie à l'est de l'Autriche, une expérimentation similaire a été conduite sur des petites et moyennes entreprises (PME) en phase de création d'activité (Gangl et al., 2014). L'échantillon est sensiblement plus faible : seules 93 entreprises sont traitées par un dispositif visant à augmenter la dissuasion de frauder qui allie l'envoi de courrier expliquant les règles en matière fiscale, des conseils pour amorcer l'activité et un audit mensuel des comptes pendant la première année de l'activité. Le groupe témoin contient 1 721 PME. La variable de consentement est mesurée par « l'opportunité du paiement », c'est-à-dire le fait pour une firme de payer l'ensemble de ces prélèvements (TVA, bénéfices, impôts locaux…) à une date fixe sans retard. Les résultats de ce dispositif sont négatifs : les PME traitées payent leurs impôts avec plus de retard que les firmes du groupe témoin.

Kosonen et Ropponen (2012) ont réalisé une expérimentation sur des PME finlandaises assujetties à la TVA, en utilisant le cadre d'une augmentation réglementaire du taux applicable passant de 9 % à 23 %. Leur traitement consistait à envoyer des courriers rédigés sous la forme de questionnaires à 300 entreprises du secteur des services tirées au sort. Les questions étaient écrites de telle sorte qu'elles faisaient apparaître les modifications des règles fiscales et interrogeaient les destinataires sur leurs connaissances de ces règles. Le groupe témoin ne recevait pas ce message. En comparant le taux de firmes qui déclaraient le bon taux de TVA dans le groupe traité et le groupe témoin, les auteurs ont trouvé que les firmes du groupe traité réalisaient significativement moins d'erreurs de déclaration, de l'ordre de 50 % de moins, que les firmes du groupe témoin. Néanmoins, le champ de l'étude semble restreint et peu généralisable à l'ensemble des entreprises finlandaises.

Enfin, l'étude la plus récente concerne une expérimentation conduite en Slovénie (Doerrenberg et Schmidt, 2014). Le protocole est identique aux précédentes études, mais les auteurs intègrent une dimension supplémentaire. En effet, un deuxième groupe traité reçoit le courrier donné en

main propre par un agent de l'administration fiscale. Ce dernier profite de sa visite pour lire le message sans apporter de précisions supplémentaires. Le message contient deux types d'informations : la première évoque l'importance de payer ses impôts parce qu'ils permettent de financer des services publics, par exemple des hôpitaux ; la seconde décrit le mécanisme d'audit des comptes ainsi que la probabilité pour une entreprise d'être contrôlée (en l'occurrence 10 %). Le secteur d'activité est restreint aux entreprises spécialisées dans la comptabilité. Le groupe témoin contient 142 entreprises. Le groupe traité qui ne reçoit pas de visite contient 32 entreprises. Le groupe traité qui reçoit une visite contient 80 entreprises. La variable retenue est le niveau de profit déclaré à l'administration fiscale. Bien que le consentement au prélèvement soit amélioré pour le groupe traité sans visite et davantage encore pour le groupe traité avec visite, ces différences ne sont pas significatives. La raison est que la taille de l'échantillon est trop faible pour détecter un effet.

Cette revue des études expérimentales destinées à améliorer le consentement au prélèvement des firmes témoigne de la diversité des protocoles expérimentaux. Cependant, les messages ne semblent pas fournir suffisamment d'incitations pour améliorer le consentement des dirigeants d'entreprise à déclarer convenablement leur base d'imposition. Nous constatons également que la portée des résultats se limite à un faible nombre de cas d'étude et qu'elle est bornée géographiquement. Enfin, il ne semble pas exister à notre connaissance d'investigations utilisant cette méthodologie qui portent sur le prélèvement sur la masse salariale pour financer la protection sociale[6].

Pour conclure, malgré les nombreux développements de la recherche sur le consentement au prélèvement, peu d'études ont tenu compte de la spécificité du comportement des firmes et de leurs réactions face aux dispositifs coercitifs et aux mécanismes moraux. Les expérimentations de terrain sont rares pour évaluer l'impact causal de dispositifs organisationnels visant à augmenter leur consentement aux prélèvements. Enfin, il n'existe pas encore d'investigations sur le consentement à la contribution sociale. Par conséquent, tous ces éléments plaident pour un enrichissement à la fois théorique et empirique. Dans ce chapitre, c'est la voie empirique que nous empruntons en conduisant une expérimentation de terrain pour améliorer le consentement à la contribution sociale des entreprises.

[6] Il existe deux études non expérimentales qui portent sur l'évaluation des dispositifs de contrôle de la fraude sociale, sur le versant fraude aux cotisations sociales (Joubert, 2009) et sur le versent fraude au travail dissimulé (Di Porto, 2011), afin de quantifier la capacité de détection des contrôles. La première étude conclut que l'amélioration de la détection permet de réduire les sommes de cotisations non prélevées. La seconde conclut que le taux de contrôle réduit le nombre de travailleurs non déclarés. Malgré leur correspondance directe avec la fraude aux prélèvements sociaux, ces deux études n'ont pas recours à une méthode expérimentale et elles se focalisent sur un unique paramètre du consentement au prélèvement.

III. Hypothèses de recherche

Dans ce chapitre, nous souhaitons tester à la fois si les prédictions du modèle canonique de la fraude fiscale peuvent s'appliquer au cadre de la contribution sociale, si celles du modèle enrichi d'interactions sociales permettent d'apprécier le comportement de déclaration sociale, si les spécificités de la contribution sociale modifient les réactions comportementales des contribuables et si la moralité contributive modifie le comportement de consentement au prélèvement social. Nous souhaitons tester ces effets grâce à l'envoi d'un message d'information dont le contenu renseigne les aspects qui rentrent dans la décision de déclarer la masse salariale. La modification de l'ensemble des informations disponibles doit permettre de faire apparaître les arbitrages des cotisants et leur décision de déclaration.

Dans un premier temps, il est possible de considérer que la contribution sociale est similaire à la fiscalité : la contribution sociale repose sur une décision volontaire de déclarer tout ou partie de sa masse salariale, sur laquelle est prélevée une cotisation à un taux donné ; l'agence de recouvrement des cotisations sociales dispose d'un outil de vérification des déclarations et peut activer des sanctions financières ou non. Dans ce cas, il est souhaitable de déterminer dans quelle mesure les paramètres de la contribution sociale seraient des déterminants du consentement au prélèvement social ? Cette question nous permet de formuler les hypothèses suivantes :

- Hypothèse 1 : l'envoi d'information sur la capacité de l'agence de recouvrement des cotisations sociales à vérifier l'exactitude des déclarations fournies par les entreprises réduit la fraude sociale ;
- Hypothèse 2 : l'envoi d'information sur le mode de redressement des cotisations non payées réduit la fraude sociale ;
- Hypothèse 3 : l'envoi d'information sur les peines d'emprisonnement consécutives à la détection pour fraude sociale réduit la fraude sociale ;
- Hypothèse 4 : l'envoi d'information sur l'existence de paiement complémentaire à l'amende, par exemple le paiement d'indemnité salariale, réduit la fraude sociale ;
- Hypothèse 5 : l'envoi d'information sur l'existence de coûts différés, par exemple l'inaccessibilité aux marchés publics, réduit la fraude sociale ;

Cependant, l'objectif de la contribution sociale est d'ouvrir un droit à prestation, tandis qu'un impôt finance le budget de l'État. Par conséquent, les deux systèmes n'ont pas la même finalité. De plus, la fraude sociale engendre une mise en danger des travailleurs illégaux, s'ils ne sont pas déclarés et des travailleurs légaux, si le niveau de cotisation payé est inférieur à celui qui leur assurerait un niveau de prestations suffisant. Cette mise en danger s'exprime pour ces travailleurs par l'absence de l'ensemble des droits auxquels ils peuvent prétendre dans un système de

protection sociale complet, comme la retraite, le chômage ou le remboursement des frais de santé. Enfin, l'acteur principal du financement de la protection sociale est l'entreprise qui a un comportement différent de celui du ménage. Il est légitime de penser que cette spécificité interagit avec la décision de se conformer aux règles de contribution sociale, notamment sur la demande de travail illégal et sur le niveau de déclaration de cotisation à payer. Dans ce cas, il est souhaitable de déterminer dans quelle mesure des paramètres spécifiques de la contribution sociale seraient des déterminants du consentement au prélèvement social ? Nous formulons les hypothèses suivantes :

- Hypothèse 6 : l'envoi d'information sur l'utilité des cotisations sociales pour financer la protection sociale collective réduit la fraude sociale ;
- Hypothèse 7 : l'envoi d'information sur le fait que l'employeur met en danger la vie d'autrui en réalisant de fausses déclarations sociales réduit la fraude sociale.

Enfin, il n'est pas déraisonnable de penser qu'à l'instar de la moralité fiscale, le système de contribution sociale contienne une moralité contributive, qui serait définie par la motivation intrinsèque à contribuer au bien-être d'autrui, afin de le protéger, selon le principe de solidarité que nous avons précédemment évoqué. Si un agent souhaite participer à la société et pouvoir en utiliser les ressources, il doit être en mesure d'y contribuer (Bourgeois, 2008). La contribution est la contrepartie de l'utilisation de la société, en l'occurrence le financement de la sécurité sociale est la contrepartie de l'utilisation de la main-d'œuvre. C'est le socle d'un compromis social forgé par la culture et les valeurs des groupes sociaux. Par conséquent, dans quelle mesure cette moralité contributive serait-elle un déterminant du consentement au prélèvement social ? Nous formulons les hypothèses suivantes :

- Hypothèse 8 : l'envoi d'information sur l'assimilation de la fraude au vol réduit la fraude sociale.
- Hypothèse 9 : l'envoi d'information sur l'idée que tout le monde contribue au financement de la protection sociale, c'est-à-dire que c'est une norme sociale, réduit la fraude sociale.
- Hypothèse 10 : l'envoi d'information sur le fait que le nom des fraudeurs peut être rendu public réduit la fraude sociale.

L'ensemble de ces tests est réalisé par le biais d'un envoi d'informations ciblées à destination des entreprises, afin d'être en mesure d'apprécier l'impact des différents paramètres du consentement au prélèvement dans les décisions déclaratives des firmes.

IV. Terrain et méthodologie

Nous avons mené une expérimentation de terrain en partenariat avec l'agence française de recouvrement des cotisations sociales. Créée en 1967, l'Acoss assure la gestion commune et centralisée des ressources et de la trésorerie de chacune des branches du régime général de la Sécurité sociale (maladie, famille, vieillesse, accidents du travail, maladies professionnelles). Elle dispose d'une autonomie financière, mais se trouve sous la double tutelle du ministère des Affaires sociales et de la Santé et du ministère de l'Économie et des Finances. La mission de l'Acoss est de garantir les délais les plus brefs entre l'encaissement des cotisations et la mise à disposition des fonds auprès des caisses prestataires, afin d'éviter des avances de trésorerie.

En 2018, l'Acoss assure le recouvrement pour le compte de 900 organisations de protection sociale (mutuelle, prévoyance, assurance). L'agence assure également le pilotage et l'animation des organismes décentralisés de recouvrement : les vingt-deux Urssaf, les quatre Caisses générales de Sécurité sociale dans les départements d'outre-mer (Cgss), et les huit centres nationaux dédiés aux offres de services. Ce pilotage concerne notamment la définition des orientations du recouvrement et du contrôle des contributions sociales, la mise en application, par les Urssaf, des textes réglementaires et législatifs, ainsi que la démarche de qualité de service et de prévention des difficultés des cotisants. Ce rôle concerne également la gestion des organismes, qu'il s'agisse des moyens budgétaires qui leur sont alloués, de la mise en œuvre de la politique informatique ou de la coordination de la politique immobilière.

1. Traitement expérimental

Le dispositif expérimental prend la forme d'un envoi de messages ciblés sur la déclaration sociale des entreprises du secteur marchand en France. Afin de réduire le coût d'envoi des informations ciblées pour contacter les entreprises, le choix a été fait d'utiliser leurs adresses électroniques. Chaque message était envoyé aux entreprises grâce à l'outil automatique d'envoi de courriels de l'Acoss via les opérateurs régionaux, les Urssaf.

Ces courriels ont respecté une apparente authenticité afin de ne pas risquer de détection de l'expérimentation, c'est-à-dire qu'ils ne contenaient pas d'informations fausses et qu'ils respectaient la mise en forme des messages habituels de l'organisme de recouvrement des cotisations sociales.

Pour tester les différentes réactions comportementales à l'envoi d'informations ciblées, nous avons constitué dix groupes traités et un groupe témoin (voir annexe n°1). Chaque groupe d'entreprises cotisantes a reçu un courriel dont le contenu correspondait à une information ciblée visant à réduire l'opportunité de frauder. Le comportement moyen de chacun des groupes est comparé à celui du comportement du groupe témoin. De cette manière, il est possible de

déterminer quels sont les facteurs comportementaux qui peuvent interagir dans le processus de déclaration du revenu à l'administration fiscale. Le groupe témoin est également destinataire d'un courriel, d'apparence « neutre », afin d'isoler l'effet propre au fait de recevoir un message, ce qui a pu constituer un biais de sélection pour des études antérieures (Fellner et al., 2013).

La campagne d'informations ciblées sur les groupes traités et le groupe témoin a eu lieu le 16 et le 26 décembre 2013, en France métropolitaine et en outre-mer, sur l'ensemble des secteurs d'activité.

Afin d'éviter une détection du protocole expérimental, nous avons restreint les établissements ciblés par la campagne d'information au niveau de leur taille. Ainsi, chaque établissement sélectionné devait avoir une taille inférieure à 250 salariés, ce qui correspond à l'ensemble des microentreprises et des PME. En 2012, ces dernières représentaient 99,9 % des entreprises en France, mais seulement 49,2 % des effectifs salariés en équivalent temps plein (Baktavatsalou et al., 2015).

2. Procédure de tirage au sort

Pour mesurer l'impact causal du dispositif expérimental sur le comportement de déclaration sociale des établissements destinataires des courriels, nous avons eu recours à une assignation aléatoire.

La base de tirage était constituée des 1 650 809 établissements ayant une taille inférieure à 250 salariés, ayant fait une déclaration dématérialisée depuis septembre 2012, c'est-à-dire pour lesquels l'Acoss disposait d'une adresse mail et ayant déclaré une masse salariale en 2013. Ces établissements ont été assignés par tirage aléatoire simple entre les différents groupes de traitement et le groupe témoin.

L'échantillonnage a également été réalisé par un tirage aléatoire simple. La moitié de l'échantillon de chaque groupe devait être composé de l'ensemble des secteurs d'activité tandis que l'autre moitié devait contenir des entreprises du secteur de la construction et du secteur de l'hébergement-restauration parce que ces deux secteurs font l'objet d'une vigilance constante de l'agence de recouvrement des cotisations sociales (Acoss, 2014). Nous avons ainsi obtenu un taux de sondage de 17,3 % des établissements de l'ensemble des secteurs d'activité, tandis que 58,5 % des établissements des secteurs de la construction et de l'hôtellerie-restauration ont fait l'objet de la campagne d'information. La taille de chaque groupe est de 40 000 établissements.

3. Sources d'information

Nous identifions chaque établissement à partir de la liste des comptes inscrits au registre de l'Acoss, disponible au sein de la base SEQUOIA. Cette base d'informations centralise depuis

janvier 1997 des extractions mensuelles des entreprises abonnées à l'Acoss qui déclarent leurs informations comptables dans les bordereaux de cotisations sociales ainsi que des données administratives. La base SEQUOIA comporte environ 600 000 comptes abonnés actifs mensuels et 1 300 000 comptes abonnés actifs trimestriels. Ainsi, les comptes sont l'unité de suivi du recouvrement des cotisations sociales et ils permettent de se rapprocher de la notion d'établissement.

Les informations administratives contenues dans la base SEQUOIA sont issues de rapprochements avec la source SIRENE[7] et elles permettent d'identifier le numéro issu du système informatique pour le répertoire des entreprises sur le territoire (Siret)[8], la raison sociale, c'est-à-dire le nom par lequel est désignée une société civile, le secteur d'activité sous la classification « activité principale exercée » (APE) et la commune de l'établissement.

Les données de comptabilité inscrites dans la base SEQUOIA sont le nombre de salariés ayant perçu des salaires au cours de la période, la date de versement des salaires, le nombre de salariés concernés par une exonération de cotisations, le montant total et le montant plafonné des rémunérations soumises à cotisations et contributions, et le montant des cotisations dues en appliquant les taux de cotisations en vigueur. Ces déclarations sociales sont situées au niveau de chaque établissement et remises aux organismes de recouvrement grâce aux bordereaux récapitulatifs de cotisations (BRC)[9]. Ils font l'objet d'un contrôle a posteriori parce qu'il existe un décalage entre la déclaration de la masse salariale et le paiement des cotisations. En effet, les établissements sont tenus de déclarer chaque mois leur masse salariale à l'organisme de recouvrement, mais ils choisissent de payer les cotisations trimestriellement ou annuellement. C'est de cette manière que l'organisme de recouvrement est en mesure de vérifier la cohérence de l'évolution de la masse salariale et d'opérer des redressements en tenant compte des déclarations antérieures. De plus, les observations de déclaration sociale sont enregistrées avec un trimestre de décalage : une déclaration au trimestre t de l'année n correspond à la situation de l'établissement du trimestre $t-1$ de l'année n. Par conséquent, nous avons recueilli des informations sur le comportement de déclarations sociales à une périodicité trimestrielle d'une période allant de 2012 à 2014.

[7] Le Système informatique du répertoire des entreprises, ou numéro SIREN1 est un code Insee unique qui sert à identifier une entreprise française. L'Insee attribue un identifiant à toute personne juridique, physique ou morale, introduite au sein d'un répertoire géré par l'Insee : SIRENE. Il est national, invariable et dure le temps de la vie de l'entreprise. Il correspond donc au numéro d'identification au répertoire (NIR) des personnes physiques.

[8] Le numéro Siret est un identifiant d'établissement. Cet identifiant numérique de 14 chiffres est articulé en deux parties : la première est le numéro Siren de l'unité légale à laquelle appartient l'unité Siret ; la seconde, habituellement appelée Nic (Numéro Interne de Classement), se compose d'un numéro d'ordre à quatre chiffres attribués à l'établissement et d'un chiffre de contrôle, qui permet de vérifier la validité de l'ensemble du numéro Siret.

[9] Le Bordereau Récapitulatif de Cotisations (BRC) est rempli par les établissements employeurs du régime général exerçant leur activité en France. Cette déclaration est mensuelle si l'effectif de l'entreprise est supérieur ou égal à 10 salariés et trimestrielle en deçà de ce seuil. Les entreprises qui déclarent mensuellement leurs cotisations représentent plus de 80 % de l'emploi total et plus de 85 % de la masse salariale totale (Acoss, 2014).

Nous utilisons également les informations issues des formulaires de Déclaration Préalable à l'Embauche (DPAE). En effet, toutes les organisations sont tenues de déclarer toute embauche de salarié relevant du régime général de la Sécurité sociale, quels que soient leur forme juridique, leur secteur d'activité, la taille de leur effectif, les conditions d'exercice de la profession et la durée d'engagement du salarié. La DPAE doit obligatoirement être effectuée, au plus tôt 8 jours avant l'embauche (et avant l'activité effective du salarié). Par ailleurs, la DPAE permet d'effectuer en une seule fois, plusieurs formalités administratives, dont l'immatriculation à l'assurance santé, à l'assurance chômage et l'examen médical d'embauche. Le non-respect de cette obligation constitue une infraction de travail dissimulé et entraîne une pénalité forfaitaire de 1 071 € par salarié concerné. C'est un indicateur utile au regard des obligations administratives que doit remplir l'employeur.

Ces déclarations sont des indicateurs de stock, c'est-à-dire ils permettent de donner une photographie chaque trimestre de la situation de chaque établissement. Cependant, ces informations ne permettent pas de saisir la dynamique de l'emploi au sein des organisations, c'est-à-dire quels sont les flux d'entrées et de sorties pendant chaque trimestre. C'est pourquoi nous construisons un ensemble d'indicateurs supplémentaires de flux d'emploi en comptabilisant le nombre de salariés déclarés sur une année. Ces indicateurs permettent de saisir d'une façon différente la dynamique du travail dissimulé, c'est-à-dire que nous souhaitons capter si les établissements ont réagi aux messages en déclarant des travailleurs présents dans l'entreprise qui étaient en situation illégale avant l'envoi des courriels.

Afin de mesurer des indicateurs de flux d'emploi, nous recourrons aux variables « postes » issues des déclarations annuelles de données sociales (DADS). Ce fichier administratif produit par l'Institut national de la statistique et des études économiques (Insee) contient d'autres formalités déclaratives que doit accomplir annuellement toute entreprise employant des salariés et pour chacun de ses établissements. Pour chaque salarié, les informations suivantes sont déclarées : la nature de l'emploi et la qualification, les dates de début et de fin de période de paie, le nombre d'heures salariées, les conditions d'emploi (temps complet, temps partiel), le montant des rémunérations versées, etc. Les DADS permettent ainsi de produire des statistiques sur les salaires et l'emploi complémentaires à celles accessibles via le fichier SEQUOIA de l'Acoss. Néanmoins, seules des variables d'entrées dans les organisations peuvent être construites. En effet, les employeurs n'ont pas l'obligation de renseigner les sorties.

4. Limites et retraitement de l'échantillon

Malgré les précautions du protocole expérimental, nous n'avons pas anticipé un phénomène particulier. Lorsque les courriels ont été envoyés aux établissements sélectionnés, ils ont été envoyés aux comptes abonnés à l'Acoss. Or, la gestion de l'abonnement peut être sous-traitée à un abonné tiers : c'est le cas d'une organisation qui ne dispose pas d'un département de comptabilité ou de ressources humaines et qui externalise la déclaration sociale, par exemple auprès d'un comptable indépendant. Ainsi, les courriels ont pu être réceptionnés par l'abonné direct ou l'abonné tiers. Nous nous retrouvons donc face à un problème de multitraitement : un établissement abonné a potentiellement pu recevoir plusieurs courriels avec un intitulé différent s'il a externalisé son service de déclaration sociale.

Nous avons également souhaité conserver les établissements pour lesquels l'historique des déclarations sociales était convenablement renseigné. En effet, pour certaines périodes d'observation, des informations étaient manquantes pour certains établissements, à cause des mises à jour du fichier SEQUOIA. Certains établissements demeurent dans la base de données sans avoir de salarié ni de rémunération à verser. Nous avons donc dû procéder à un retraitement de l'échantillon.

Notre stratégie d'évaluation d'impact repose donc sur une sélection d'un échantillon de nos groupes traités et de notre groupe témoin, par un filtrage successif des établissements concernés par l'expérimentation : un premier filtre qui conserve les établissements cotisants directement abonnés à l'Acoss et n'étant pas concerné par un abonnement tiers ; un deuxième filtre permet de conserver les établissements qui ne changent pas d'identité administrative au cours de l'analyse (changement de SIRET, changement de localisation, changement de raison sociale, changement de secteur d'activité) ; un troisième filtre permet de conserver les établissements dont la taille n'est pas supérieure à 250 employés en 2013 et 2014 ; un quatrième filtre permet de conserver les établissements actifs en 2013 et 2014, c'est-à-dire ceux qui n'ont pas connu de cessation d'activité après l'envoi des courriels.

V. Données

1. Participants

Les principales caractéristiques de cet échantillon d'établissements servant à l'analyse empirique sont présentées dans le tableau 1. Les tests d'équivalence entre le groupe témoin et les groupes traités figurent également dans ce tableau.

Parmi les 77 476 établissements sélectionnés, un peu plus du quart des établissements est localisé dans la région Île-de-France. Environ 66 % des établissements sont des sociétés commerciales. Le secteur d'activité modal regroupe les établissements du secteur de la construction (25 %) et le deuxième secteur d'activité le plus représenté est celui de l'hébergement-restauration (13 %), conformément à la décomposition issue du protocole expérimental.

Par rapport au groupe témoin, les différences que nous observons à l'issue du tirage au sort et du retraitement décrit dans la partie IV sont les suivantes :

- le groupe ayant reçu le courriel insistant sur le risque de détection contient moins d'établissements localisés en Champagne-Ardenne et moins de personnes morales soumises au droit étranger ;
- le groupe ayant reçu le courriel présentant les sanctions monétaires possède plus d'établissements situés en Bourgogne ;
- le groupe ayant reçu le courriel décrivant les sanctions non monétaires contient davantage d'établissements localisés dans les départements d'outre-mer et dans le Nord–Pas-de-Calais, mais moins d'établissements situés en Île-de-France, et il y a moins de personnes morales de droit étranger ainsi que d'établissements dans le secteur de l'industrie extractive ;
- le groupe ayant reçu le courriel insistant sur les coûts dissimulés de la fraude possède moins d'établissements localisés en Champagne-Ardenne, moins de personnalités morales soumises au droit administratif, moins d'administrations publiques et moins d'établissements du secteur de l'enseignement ;
- le groupe ayant reçu le courriel décrivant les coûts différés contient plus d'établissements situés en Franche-Comté, et moins d'entreprises d'industries extractives ;
- le groupe ayant reçu le courriel insistant sur la contribution à un bien public a davantage d'établissements issus des régions du Nord–Pas-de-Calais et de Franche-Comté ;
- le groupe ayant reçu un courriel contenant l'idée que la fraude c'est du vol a plus d'établissements localisés dans les départements d'outre-mer, moins de personnes morales soumises au droit étranger et moins d'établissements relevant du secteur de l'information-communication ;
- le groupe ayant reçu le courriel qui évoquait la possibilité de rendre public le nom des fraudeurs contient plus d'établissements situés en Basse-Normandie et moins d'établissements du secteur de l'enseignement ;
- le groupe ayant reçu le courriel insistant sur la mise en danger d'autrui possède plus d'établissements de Franche-Comté et du Nord–Pas-de-Calais, moins d'entreprises de sylviculture et moins d'industries extractives ;

- le groupe ayant reçu le courriel contenant l'idée que le prélèvement est la norme sociale contient plus d'établissements en Nord–Pas-de-Calais, moins d'établissements en Midi-Pyrénées, tandis qu'il possède plus d'établissements relevant du secteur de l'extraction et des activités spécialisées, scientifiques et techniques, et plus de personnalités morales « autres » inscrites au registre des commerces et société (RCS) (par exemple société coopérative agricole).

Tableau 1 : Caractéristiques des établissements traités et témoin

	Témoin	Risque de détection	Sanctions monétaires	Sanctions non monétaires	Coûts cachés	Coûts différés	Contribution à la production d'un bien public	« Frauder c'est voler »	Publicité du nom des fraudeurs	Mise en danger d'autrui	Norme sociale
Nombre d'établissements	6898	7031	7052	7048	7125	7182	6943	7025	7039	7075	7058
Alsace	2 609*** (0,200)	0,164 (0,281)	0,113 (0,281)	0,313 (0,281)	0,001 07 (0,280)	0,426 (0,279)	0,156 (0,282)	0,309 (0,281)	0,402 (0,281)	0,147 (0,281)	0,366 (0,281)
Aquitaine	5 001*** (0,264)	0,019 2 (0,372)	0,089 3 (0,372)	-0,305 (0,372)	-0,244 (0,371)	0,484 (0,370)	0,500 (0,373)	-0,076 2 (0,372)	-0,256 (0,372)	0,228 (0,371)	0,340 (0,372)
Auvergne	1 725*** (0,153)	-0,061 1 (0,215)	-0,222 (0,215)	-0,093 5 (0,215)	0,085 4 (0,214)	-0,249 (0,214)	-0,184 (0,215)	0,040 0 (0,215)	-0,077 2 (0,215)	-0,043 2 (0,214)	-0,209 (0,215)
Basse-Normandie	1 885*** (0,171)	0,192 (0,241)	0,143 (0,241)	**0,456*** **(0,241)**	0,375 (0,241)	0,287 (0,240)	0,059 8 (0,242)	0,151 (0,241)	**0,431*** **(0,241)**	0,009 39 (0,241)	-0,071 1 (0,241)
Bourgogne	2 696*** (0,190)	-0,036 8 (0,267)	**-0,612**** **(0,267)**	-0,384 (0,267)	0,139 (0,266)	-0,050 9 (0,266)	-0,205 (0,268)	-0,106 (0,267)	-0,310 (0,267)	-0,166 (0,267)	0,151 (0,267)
Bretagne	4 016*** (0,243)	0,009 38 (0,342)	0,295 (0,342)	0,156 (0,342)	0,251 (0,341)	0,022 2 (0,340)	0,464 (0,343)	0,369 (0,342)	0,218 (0,342)	0,394 (0,342)	0,518 (0,342)
Centre-Val de Loire	3 813*** (0,233)	0,483 (0,328)	0,285 (0,328)	-0,251 (0,328)	0,215 (0,327)	0,169 (0,326)	0,177 (0,329)	-0,083 2 (0,328)	-0,162 (0,328)	0,159 (0,328)	-0,043 9 (0,328)
Champagne-Ardenne	2 479*** (0,176)	**-0,431*** **(0,248)**	-0,224 (0,248)	-0,322 (0,248)	**-0,795***** **(0,247)**	-0,390 (0,246)	-0,275 (0,248)	-0,273 (0,248)	-0,063 9 (0,248)	-0,147 (0,247)	-0,311 (0,247)
Corse	0,623*** (0,091 3)	-0,011 8 (0,128)	-0,070 3 (0,128)	0,071 9 (0,128)	-0,076 0 (0,128)	-0,024 7 (0,128)	-0,076 1 (0,129)	0,002 97 (0,129)	-0,026 7 (0,128)	-0,100 (0,128)	-0,184 (0,128)
DOM	3 189*** (0,214)	-0,160 (0,301)	0,101 (0,301)	**0,528*** **(0,301)**	-0,256 (0,300)	-0,056 5 (0,300)	0,195 (0,302)	**0,597**** **(0,301)**	0,277 (0,301)	-0,108 (0,301)	-0,256 (0,301)
Franche-Comté	1 290*** (0,145)	0,175 (0,205)	0,185 (0,205)	0,270 (0,205)	0,029 1 (0,204)	**0,339*** **(0,204)**	**0,366*** **(0,205)**	-0,109 (0,205)	0,173 (0,205)	**0,378*** **(0,204)**	0,297 (0,205)
Haute Normandie	2 754*** (0,194)	0,147 (0,273)	0,110 (0,273)	0,168 (0,273)	-0,411 (0,273)	-0,290 (0,272)	-0,046 7 (0,274)	0,206 (0,274)	-0,197 (0,273)	-0,097 2 (0,273)	-0,445 (0,273)
Île-de-France	26,04*** (0,526)	-0,151 (0,741)	-0,441 (0,740)	**-1 334*** **(0,740)**	-0,380 (0,738)	0,460 (0,737)	-0,932 (0,743)	0,170 (0,741)	-0,692 (0,741)	-0,425 (0,740)	0,274 (0,740)
Languedoc-Roussillon	3 117*** (0,214)	0,054 8 (0,301)	0,017 0 (0,301)	0,459 (0,301)	0,420 (0,300)	-0,109 (0,300)	-0,005 80 (0,302)	0,385 (0,301)	0,179 (0,301)	-0,063 8 (0,301)	0,312 (0,301)

	Témoin	Risque de détection	Sanctions monétaires	Sanctions non monétaires	Coûts cachés	Coûts différés	Contribution à la production d'un bien public	« Frauder c'est voler »	Publicité du nom des fraudeurs	Mise en danger d'autrui	Norme sociale
Limousin	1 116*** (0,122)	-0,021 1 (0,172)	-0,095 3 (0,172)	0,033 0 (0,172)	-0,204 (0,172)	-0,099 8 (0,171)	-0,050 4 (0,173)	-0,205 (0,172)	0,034 5 (0,172)	-0,070 3 (0,172)	-0,124 (0,172)
Lorraine	3 102*** (0,211)	0,140 (0,296)	0,301 (0,296)	0,075 9 (0,296)	-0,070 8 (0,295)	-0,011 3 (0,295)	0,167 (0,297)	**-0,526*** **(0,296)**	0,364 (0,296)	-0,120 (0,296)	0,298 (0,296)
Midi-Pyrénées	4 146*** (0,235)	-0,221 (0,331)	0,093 8 (0,331)	-0,372 (0,331)	0,275 (0,330)	-0,331 (0,329)	-0,329 (0,332)	-0,146 (0,331)	-0,211 (0,331)	0,193 (0,331)	**-0,845**** **(0,331)**
Nord Pas de Calais	3 247*** (0,227)	0,493 (0,320)	0,440 (0,320)	**0,626*** **(0,320)**	0,318 (0,319)	0,415 (0,318)	**0,699**** **(0,321)**	0,511 (0,320)	0,304 (0,320)	**0,640**** **(0,319)**	**0,550*** **(0,320)**
Pays de la Loire	5 277*** (0,266)	-0,142 (0,374)	-0,172 (0,374)	0,171 (0,374)	-0,041 8 (0,373)	-0,334 (0,372)	-0,0774 (0,375)	-0,0954 (0,374)	-0,00626 (0,374)	-0,443 (0,374)	-0,346 (0,374)
Picardie	2,696*** (0,192)	-0,435 (0,270)	-0,101 (0,270)	-0,000634 (0,270)	0,0123 (0,269)	0,0883 (0,268)	-0,0751 (0,271)	-0,262 (0,270)	0,102 (0,270)	0,0173 (0,269)	-0,373 (0,270)
Poitou-Charentes	2,667*** (0,194)	-0,136 (0,273)	0,183 (0,273)	0,0142 (0,273)	0,238 (0,272)	-0,300 (0,271)	0,184 (0,274)	-0,0767 (0,273)	0,0318 (0,273)	-0,208 (0,272)	0,0245 (0,273)
PACA	7,147*** (0,306)	0,0781 (0,431)	-0,355 (0,431)	-0,138 (0,431)	-0,0312 (0,429)	-0,0181 (0,429)	-0,680 (0,432)	-0,613 (0,431)	-0,115 (0,431)	0,0473 (0,430)	-0,389 (0,430)
Rhône-Alpes	9,278*** (0,348)	-0,0902 (0,489)	-0,0608 (0,489)	-0,127 (0,489)	0,182 (0,488)	-0,478 (0,487)	-0,0745 (0,491)	-0,196 (0,490)	-0,427 (0,489)	-0,246 (0,489)	0,470 (0,489)
Étranger	0,0870** (0,0367)	-0,0585 (0,0516)	-0,00190 (0,0516)	-0,0160 (0,0516)	-0,0308 (0,0515)	0,0523 (0,0514)	0,0426 (0,0518)	0,0269 (0,0516)	0,0267 (0,0516)	0,0261 (0,0516)	-0,00197 (0,0516)
Entrepreneur individuel	8,075*** (0,332)	0,0322 (0,467)	0,604 (0,467)	0,410 (0,467)	0,346 (0,466)	-0,124 (0,465)	0,538 (0,469)	0,210 (0,467)	0,435 (0,467)	-0,414 (0,467)	0,370 (0,467)
Groupement de droit privé non doté de la personnalité morale	0,333*** (0,0703)	-0,0348 (0,0989)	-0,00728 (0,0989)	0,00709 (0,0989)	0,0736 (0,0986)	0,0147 (0,0984)	-0,0886 (0,0993)	-0,0772 (0,0990)	0,121 (0,0989)	0,0199 (0,0988)	0,0633 (0,0989)
Personne morale de droit étranger	0,333*** (0,0611)	**-0,163*** **(0,0860)**	-0,0215 (0,0859)	**-0,192**** **(0,0859)**	-0,123 (0,0857)	0,000739 (0,0855)	-0,103 (0,0863)	**-0,205**** **(0,0860)**	-0,0919 (0,0860)	-0,00834 (0,0859)	0,0775 (0,0859)
Personne morale de droit public soumise au droit commercial	0,203*** (0,0536)	0,0246 (0,0755)	-0,0470 (0,0754)	0,0666 (0,0754)	-0,0486 (0,0752)	0,0477 (0,0751)	0,0419 (0,0757)	-0,0748 (0,0755)	-0,0325 (0,0755)	-0,104 (0,0754)	0,0804 (0,0754)

	Témoin	Risque de détection	Sanctions monétaires	Sanctions non monétaires	Coûts cachés	Coûts différés	Contribution à la production d'un bien public	« Frauder c'est voler »	Publicité du nom des fraudeurs	Mise en danger d'autrui	Norme sociale
Société commerciale	66,79*** (0,569)	-0,140 (0,801)	-1,147 (0,800)	-0,783 (0,800)	0,118 (0,798)	-0,942 (0,797)	-0,562 (0,803)	-0,382 (0,801)	-1,210 (0,801)	0,817 (0,800)	-0,976 (0,800)
Autre personne morale immatriculée au registre du commerce et des sociétés (RCS)	1,682*** (0,162)	0,281 (0,228)	0,148 (0,228)	0,276 (0,228)	0,269 (0,227)	-0,0943 (0,227)	0,349 (0,229)	0,0835 (0,228)	0,0373 (0,228)	0,000332 (0,228)	**0,472**** **(0,228)**
Personne morale et organisme soumis au droit administratif	10,76*** (0,371)	-0,431 (0,522)	0,304 (0,522)	-0,257 (0,522)	**-0,890*** **(0,521)**	0,424 (0,520)	-0,0553 (0,524)	-0,337 (0,523)	0,324 (0,522)	-0,0995 (0,522)	-0,329 (0,522)
Organisme privé spécialisé	0,580*** (0,0890)	-0,0963 (0,125)	-0,0836 (0,125)	0,0160 (0,125)	-0,00444 (0,125)	0,0328 (0,125)	-0,119 (0,126)	-0,0532 (0,125)	0,145 (0,125)	-0,000373 (0,125)	-0,169 (0,125)
Groupement de droit privé	11,25*** (0,385)	0,527 (0,542)	0,251 (0,542)	0,456 (0,542)	0,259 (0,540)	0,641 (0,539)	-0,000898 (0,544)	0,836 (0,542)	0,272 (0,542)	-0,211 (0,541)	0,411 (0,541)
Agriculture sylviculture et pêche	0,145*** (0,0410)	-0,0596 (0,0577)	0,0110 (0,0577)	-0,00309 (0,0577)	-0,0467 (0,0575)	-0,0336 (0,0574)	-0,000940 (0,0579)	-0,0169 (0,0577)	-0,0313 (0,0577)	**-0,103*** **(0,0576)**	-0,0316 (0,0577)
Industries extractives	0,203*** (0,0421)	-0,0750 (0,0593)	-0,0753 (0,0593)	**-0,132**** **(0,0593)**	-0,0205 (0,0591)	**-0,147**** **(0,0590)**	-0,0445 (0,0595)	-0,0891 (0,0593)	-0,0609 (0,0593)	**-0,146**** **(0,0592)**	-0,0896 (0,0592)
Industrie manufacturière	5,494*** (0,277)	0,365 (0,390)	-0,0207 (0,389)	-0,159 (0,390)	0,597 (0,389)	-0,0920 (0,388)	-0,280 (0,391)	0,342 (0,390)	-0,210 (0,390)	0,555 (0,389)	0,102 (0,389)
Production et distribution d'électricité, de gaz, de vapeur et d'air conditionné	0,0435 (0,0296)	0,0418 (0,0417)	0,0416 (0,0417)	-0,000926 (0,0417)	0,0267 (0,0416)	-0,0156 (0,0415)	0,0285 (0,0419)	0,0277 (0,0417)	0,0133 (0,0417)	0,0272 (0,0417)	-0,000986 (0,0417)
Production et distribution d'eau ; assainissement, gestion des déchets et dépollution	0,652*** (0,0970)	0,0446 (0,137)	-0,0284 (0,136)	0,0571 (0,136)	0,00729 (0,136)	-0,0119 (0,136)	-0,00423 (0,137)	0,0309 (0,137)	-0,0699 (0,136)	0,0544 (0,136)	-0,0715 (0,136)
Construction	25,56*** (0,524)	-0,128 (0,738)	0,746 (0,737)	-1,168 (0,737)	0,238 (0,735)	-0,301 (0,734)	-0,122 (0,740)	-0,604 (0,738)	-0,0716 (0,738)	0,604 (0,737)	-0,721 (0,737)
Commerce ; réparation d'automobiles et de motocycles	12,42*** (0,392)	-0,676 (0,552)	-0,399 (0,552)	-0,307 (0,552)	-0,466 (0,550)	-0,477 (0,549)	0,481 (0,554)	-0,253 (0,552)	-0,874 (0,552)	-0,452 (0,551)	-0,480 (0,552)
Transports et entreposage	2,276*** (0,178)	-0,143 (0,250)	-0,135 (0,250)	0,108 (0,250)	-0,129 (0,250)	-0,0204 (0,249)	-0,288 (0,251)	0,172 (0,250)	0,196 (0,250)	-0,198 (0,250)	-0,0374 (0,250)
Hébergement et restauration	13,28*** (0,410)	-0,0379 (0,577)	-0,0915 (0,576)	0,923 (0,576)	0,110 (0,575)	0,0179 (0,574)	-0,259 (0,579)	-0,0550 (0,577)	0,544 (0,577)	-0,247 (0,576)	0,110 (0,576)

	Témoin	Risque de détection	Sanctions monétaires	Sanctions non monétaires	Coûts cachés	Coûts différés	Contribution à la production d'un bien public	« Frauder c'est voler »	Publicité du nom des fraudeurs	Mise en danger d'autrui	Norme sociale
Information et communication	2,044***	0,246	0,253	-0,0861	0,159	-0,164	0,188	**-0,407***	-0,0125	0,0619	-0,0888
	(0,171)	(0,241)	(0,240)	(0,240)	(0,240)	(0,239)	(0,241)	**(0,241)**	(0,240)	(0,240)	(0,240)
Activités financières et d'assurance	2,537***	-0,133	**-0,467***	0,116	-0,221	-0,337	0,214	-0,0886	0,120	-0,0211	-0,327
	(0,185)	(0,261)	**(0,261)**	(0,261)	(0,260)	(0,260)	(0,262)	(0,261)	(0,261)	(0,261)	(0,261)
Activités immobilières	1,392***	-0,0548	0,0972	0,141	0,152	0,168	-0,0810	0,00331	-0,0279	-0,134	0,323
	(0,144)	(0,202)	(0,202)	(0,202)	(0,202)	(0,201)	(0,203)	(0,202)	(0,202)	(0,202)	(0,202)
Activités spécialisées, scientifiques et techniques	5,335***	0,539	-0,0740	0,113	0,644	0,680*	0,0662	0,459	0,305	0,404	**1,324*** **
	(0,280)	(0,394)	(0,394)	(0,394)	(0,393)	(0,392)	(0,395)	(0,394)	(0,394)	(0,394)	**(0,394)**
Activités de services administratifs et de soutien	3,871***	0,211	0,185	0,499	0,0170	0,348	0,407	0,357	0,278	0,101	0,536
	(0,240)	(0,338)	(0,337)	(0,337)	(0,336)	(0,336)	(0,339)	(0,338)	(0,337)	(0,337)	(0,337)
Administration publique	8,234***	-0,426	0,0471	-0,118	**-0,852***	0,663	0,0619	-0,206	0,645	-0,0647	-0,328
	(0,330)	(0,464)	(0,464)	(0,464)	**(0,463)**	(0,462)	(0,466)	(0,465)	(0,464)	(0,464)	(0,464)
Enseignement	2,537***	-0,176	0,101	-0,0398	**-0,460***	-0,198	-0,305	0,0680	**-0,690*** **	-0,346	-0,0150
	(0,182)	(0,257)	(0,256)	(0,257)	**(0,256)**	(0,255)	(0,257)	(0,257)	**(0,257)**	(0,256)	(0,256)
Santé humaine et action sociale	5,480***	-0,104	0,0221	0,0111	0,148	-0,523	0,209	-0,327	-0,0387	-0,377	-0,323
	(0,271)	(0,382)	(0,381)	(0,381)	(0,380)	(0,380)	(0,383)	(0,382)	(0,382)	(0,381)	(0,381)
Arts, spectacles et activités récréatives	3,885***	0,0403	0,213	0,0876	-0,194	0,320	-0,241	0,0436	0,0642	0,0724	-0,286
	(0,233)	(0,328)	(0,328)	(0,328)	(0,327)	(0,326)	(0,329)	(0,328)	(0,328)	(0,327)	(0,328)
Autres activités de services	4,596***	0,496	-0,427	-0,0694	0,275	0,0828	-0,0154	0,543	-0,0778	0,210	0,349
	(0,255)	(0,359)	(0,359)	(0,359)	(0,358)	(0,358)	(0,361)	(0,359)	(0,359)	(0,359)	(0,359)
Activités extraterritoriales	0,0145	0,0282	-0,000317	0,0281	0,0136	0,0412	-0,0145	-0,000262	-0,000290	-0,000363	**0,0563** **
	(0,0203)	(0,0286)	(0,0285)	(0,0285)	(0,0285)	(0,0284)	(0,0286)	(0,0286)	(0,0285)	(0,0285)	**(0,0285)**

Note : le groupe témoin sert de valeur de référence. Les enregistrements des groupes traités sont les écarts à cette valeur de référence. Les erreurs-types sont entre parenthèses. *** : p-value inférieure à 0,01 ; ** : p-value inférieure à 0,05 ; * : p-value inférieure à 0,10.

2. Les indicateurs d'intérêts

Nous mesurons le consentement à la contribution sociale à travers deux jeux d'indicateurs : les déclarations sociales obligatoires et les flux d'emplois.

a. Les indicateurs de stock : les déclarations sociales permettent d'apprécier le comportement de dissimulation de l'activité

Nous observons pour chaque établissement l'effectif déclaré en fin de trimestre. Nous analysons également le nombre de déclarations d'embauche trimestrielles en contrat à durée indéterminée, à durée déterminée de plus d'un mois (CDD long) et à durée déterminée de moins d'un mois (CDD court). De cette manière, nous suivons la quantité de main-d'œuvre employée afin d'être capable de quantifier le nombre absolu de personnes qui est déclaré à la caisse de recouvrement des cotisations sociales. Cet indicateur permet de mesurer le comportement de consentement au prélèvement social sur le versant effectif.

Nous examinons également l'assiette déplafonnée des cotisations d'assurance sociale. L'ensemble des rémunérations versées dans le cadre d'un contrat de travail est inclus dans le calcul de l'assiette, qui sert de base pour l'application des taux de cotisations sociales. Dans l'assiette sont compris les salaires, les indemnités, les prestations sociales complémentaires, les revenus de remplacement, ainsi que les avantages en espèces ou en nature. L'assiette minimale retenue se calcule au niveau du Salaire minimum interprofessionnel de croissance (SMIC). L'assiette maximale correspond au plafond de cotisations pour lequel il n'y a plus de prélèvements sociaux sur les rémunérations. Ainsi, l'assiette déplafonnée correspond à l'assiette salariale totale, autrement dit c'est un indicateur de la masse salariale de l'établissement. Nous analysons également l'évolution de l'assiette Contribution Sociale Généralisée (CSG), qui se distingue de l'assiette déplafonnée parce qu'elle est constituée des éléments de rémunération non soumis à cotisation sociale : l'intéressement, la participation, des indemnités de rupture du contrat de travail (conventionnelle ou non). Ces deux indicateurs sont renseignés à un niveau agrégé et nous ne sommes pas en mesure de décomposer chaque assiette par les montants qui les composent.

b. Les indicateurs de flux d'emplois

Les flux sont l'ensemble des créations et destructions d'emplois au sein des entreprises, que l'on approche par les variations annuelles d'emploi, en distinguant les flux positifs et les flux négatifs. Autrement dit, c'est la somme des gains en effectif des entreprises dont l'emploi croît (création) et les pertes de celles dont l'emploi décroît (destruction). Traditionnellement, on mesure les flux d'emplois à travers quatre indicateurs (Davis et Haltiwanger, 1999) :

- le taux de création d'emploi, qui est le rapport entre la somme des emplois créés, C, entre deux périodes, en l'occurrence deux années identifiées par l'indice a, pour les établissements, e, pour lesquels la variation d'emploi entre deux périodes est positive et la taille moyenne entre les deux périodes, Z, qui est le nombre moyen d'individus en poste pour un contrat donné $c = \{cdi, cdd\}$ entre l'année considérée et l'année antérieure :

$$(c_{c_{poste}})_{e_a} = \frac{C_{e_a}}{Z_{e_a}}$$

$$(c_{c_{poste}})_{e_a} = \frac{\sum_{e \in e^+} \Delta(n_{c_{poste}})_{e_a}}{\frac{1}{2} \cdot \left(\left(n_{c_{poste}}\right)_{e_a} + \left(n_{c_{poste}}\right)_{e_{a-1}} \right)}$$

$$(c_{c_{poste}})_{e_a} = \frac{\sum_{e \in e^+} \left(\left(n_{c_{poste}}\right)_{e_a} - \left(n_{c_{poste}}\right)_{e_{a-1}} \right)}{\frac{1}{2} \cdot \left(\left(n_{c_{poste}}\right)_{e_a} + \left(n_{c_{poste}}\right)_{e_{a-1}} \right)}$$

- le taux de destruction d'emploi, qui est le rapport entre la somme des emplois détruits, D, en valeur absolue entre deux périodes, en l'occurrence deux années identifiées par l'indice a, pour les établissements e pour lesquels la variation d'emploi entre deux périodes est négative et la taille moyenne entre les deux périodes, Z :

$$(d_{c_{poste}})_{e_a} = \frac{D_{e_a}}{Z_{e_a}}$$

$$(d_{c_{poste}})_{e_a} = \frac{\sum_{e \in e^-} \left| \Delta(n_{c_{poste}})_{e_a} \right|}{\frac{1}{2} \cdot \left(\left(n_{c_{poste}}\right)_{e_a} + \left(n_{c_{poste}}\right)_{e_{a-1}} \right)}$$

$$(d_{c_{poste}})_{e_a} = \frac{\sum_{e \in e^-} \left| \left(n_{c_{poste}}\right)_{e_a} - \left(n_{c_{poste}}\right)_{e_{a-1}} \right|}{\frac{1}{2} \cdot \left(\left(n_{c_{poste}}\right)_{e_a} + \left(n_{c_{poste}}\right)_{e_{a-1}} \right)}$$

- le taux de croissance de l'emploi, qui est le rapport entre la variation d'emploi G entre deux années a, et la taille de l'établissement Z,

$$(g_{c_{poste}})_{e_a} = \frac{G_{e_a}}{Z_{e_a}}$$

$$(g_{c_{poste}})_{e_a} = \frac{\Delta\left(n_{c_{poste}}\right)_{e_a}}{\frac{1}{2}.\left(\left(n_{c_{poste}}\right)_{e_a} + \left(n_{c_{poste}}\right)_{e_{a-1}}\right)}$$

$$(g_{c_{poste}})_{e_a} = \frac{\left(n_{c_{poste}}\right)_{e_a} - \left(n_{c_{poste}}\right)_{e_{a-1}}}{\frac{1}{2}.\left(\left(n_{c_{poste}}\right)_{e_a} + \left(n_{c_{poste}}\right)_{e_{a-1}}\right)}$$

Généralement, les flux d'emplois sont mesurés au niveau du secteur d'activité, afin d'identifier la tension sur le marché du travail et de déterminer les secteurs où les flux sont les plus importants ou les moins importants (Picart, 2008). Nous déterminons toutes les créations et les destructions au niveau de chaque établissement de chaque groupe, qui correspondent respectivement aux variations positives de stock d'une année sur l'autre et aux variations négatives de stock d'une année sur l'autre. Puis, nous déterminons la taille de chaque établissement. Cet indicateur est sommé pour obtenir la taille de chaque groupe expérimental : c'est la variable Z au niveau d'un groupe. Analytiquement nous procédons de la façon suivante :

$$z_{e_a} = \frac{1}{2}.\left(\left(n_{c_{poste}}\right)_{e_a} + \left(n_{c_{poste}}\right)_{e_{a-1}}\right)$$

$$Z_{G_a} = \sum z_{e_a}$$

$$Z_{G_a} = \sum\left[\frac{1}{2}.\left(\left(n_{c_{poste}}\right)_{e_a} + \left(n_{c_{poste}}\right)_{e_{a-1}}\right)\right]$$

On rapporte ensuite pour chaque établissement la création et la destruction d'emplois à cette variable de taille d'établissement, et nous multiplions cette quantité par le nombre d'établissements de chaque groupe. Analytiquement, on a un indicateur de création établissement qui se compose de la manière suivante :

$$\left(y_{c_{poste}}\right)_{e_a} = \frac{\sum y_{e_a}}{Z_{G_a}}.N$$

En calculant la moyenne de cette variable, on obtient la quantité suivante :

$$\overline{\left(y_{c_{poste}}\right)_{e_a}} = \frac{1}{N}.\sum \frac{\sum y_{e_a}}{Z_{G_a}}.N$$

$$\overline{\left(y_{c_{poste}}\right)_{e_a}} = \sum \frac{\sum y_{e_a}}{Z_{G_a}}$$

On montre qu'on obtient l'indicateur de création ou de destruction d'emploi pour chacun des groupes qui a la forme suivante :

$$Y_{G_a} = \frac{\sum y_{e_a}}{\sum z_{e_a}}$$

Lorsqu'on procède de cette manière, il est possible de calculer l'écart-type de cet indicateur $\left(y_{c_{poste}}\right)_{e_a}$. C'est pour cela que nous utilisons cette méthode de calcul, qui nous servira pour pouvoir comparer les créations/destructions entre le groupe témoin et les groupes tests dans un modèle de régression. On réalise donc une comparaison d'un indicateur de création/destruction d'emploi moyen au niveau de chaque établissement qui est en réalité le taux de création/destruction de chaque groupe d'établissements.

D'une certaine façon, cet indicateur permet de tenir compte de chaque flux individuel d'établissement dans l'ensemble du stock du groupe considéré. C'est cette variabilité que l'on cherche à obtenir, en comparant les indicateurs entre chaque groupe traité et témoin.

3. Méthode d'identification des effets

En présence du problème de multitraitement, nous avons été conduits à opérer une sélection de notre échantillon. Par conséquent, l'estimation de l'effet causal est dépendante de cette sélection. Dans un premier temps, nous recourrons à une estimation non paramétrique de l'effet causal du traitement. Dans un second temps, nous estimons l'effet causal en utilisant la dimension panel de nos données, ce qui nous permet de tenir compte de ces différences.

a. L'estimation non-paramétrique de l'effet causal en utilisant la dimension coupe transversale

Le protocole expérimental, en assignant aléatoirement des établissements à la réception de différents courriels dont le contenu était destiné à faire révéler les arbitrages en matière de déclarations sociales, nous permet d'estimer les écarts observés entre les indicateurs d'intérêt pour chacun des groupes traités par rapport au groupe témoin comme l'effet causal du programme sur le comportement de déclaration sociale. Il s'agit ensuite de mesurer les écarts de moyennes à la valeur de référence du groupe témoin. Nous utilisons la statistique de Student pour réaliser le test

d'hypothèse de nullité des différences. Nous effectuons ces tests de comparaison de moyenne chaque trimestre à partir de l'envoi des courriers.

b. Estimation paramétrique de l'effet causal en utilisant la dimension panel

Néanmoins, la sélection d'un sous-échantillon à partir de l'échantillon initial conduit à un biais de sélection. Les écarts peuvent être engendrés par un déterminant différent de ceux causés par les messages envoyés pendant la campagne d'information.

En identifiant les comportements de déclaration sociale antérieurs et postérieurs à la campagne d'informations ciblées, nous pouvons ainsi tenir compte de l'ensemble de ces perturbations dans une modélisation économétrique afin d'estimer l'impact causal du traitement purgé des différences observables et inobservables avant l'expérimentation.

Ainsi, la dimension des données en panel créée à partir du suivi trimestriel des entreprises conduit notre stratégie empirique à adopter une analyse complémentaire en recourant à une différence de différences. En présence de plusieurs groupes de traitement et de plusieurs trimestres d'observation, il est possible de choisir la modélisation économétrique suivante (Wooldridge, 2010) :

$$Y_{i_{g_t}} = \sum \alpha_{1_g}.G_g + \sum \alpha_{2_t}.T_t + \sum \alpha_{3_{g_t}}.G_g.T_t + \beta + \epsilon_{i_{t_a}}$$

Nos variables d'intérêt, $Y_{i_{g_t}}$, mesurées au niveau de chaque établissement, sont les effectifs en fin de trimestre, les déclarations préalables à l'embauche et le calcul des assiettes soumises à cotisation, ainsi que flux d'emplois annuels, qui sont des données déclarées par les établissements cotisants. Le modèle tient compte d'une variable dichotomique G_g, qui permet de renseigner l'appartenance de l'établissement à un groupe traité ($G_1 = 1$) ou au groupe témoin ($G_0 = 0$) et de tenir compte d'un effet fixe « groupe d'établissements ». Il prend en compte d'une deuxième variable dichotomique T_t qui indique à quel moment intervient l'expérimentation afin de différencier la période avant le début de l'expérimentation ($T_0 = 0$) de la période après le début de l'expérimentation ($T_1 = 1$). Le terme d'interaction $G_g.T_t$ prend la forme d'une troisième variable dichotomique qui permet l'identification des établissements traités après le début de l'expérimentation.

Dans ce modèle, chaque estimation des coefficients permet d'obtenir une estimation de différents indicateurs d'intérêts :

$\beta = \overline{Y_{0_0}}$: la moyenne de l'indicateur d'intérêt de déclaration sociale pour le groupe témoin avant l'envoi des courriels ;

$\alpha_1 = \overline{Y_{1_0}} - \overline{Y_{0_0}}$: la différence d'un indicateur d'intérêt de déclaration sociale entre le groupe traité et le groupe témoin avant l'envoi des courriels ;

$\alpha_2 = \overline{Y_{0_1}} - \overline{Y_{0_0}}$: la différence d'un indicateur d'intérêt de déclaration sociale pour le groupe témoin entre avant et après l'envoi des courriels ;

$\alpha_3 = \left(\overline{Y_{1_1}} - \overline{Y_{0_1}}\right) - (\overline{Y_{1_0}} - \overline{Y_{0_0}})$: la double différence d'un indicateur d'intérêt de déclaration sociale entre le groupe traité et le groupe témoin entre avant et après l'envoi des courriels : c'est l'estimateur de l'effet causal du traitement, i.e. l'envoi d'une information destinée à désinciter les entreprises à adopter un comportement frauduleux, sur chacun des groupes destinataires de messages électroniques ;

$\alpha_1 + \alpha_3 = \overline{Y_{1_1}} - \overline{Y_{0_1}}$: la différence d'un indicateur d'intérêt de déclaration sociale entre le groupe traité et le groupe témoin après l'envoi des courriels, elle correspond à l'estimation naïve de l'intention de traiter qui ne tient donc pas compte de la période avant l'envoi du courriel ;

$\alpha_2 + \alpha_3 = \overline{Y_{1_1}} - \overline{Y_{1_0}}$: la différence d'un indicateur d'intérêt de déclaration sociale pour le groupe traité entre avant et après l'envoi des courriels ;

$\alpha_2 - \alpha_1 = \overline{Y_{0_1}} - \overline{Y_{1_0}}$: la différence d'un indicateur d'intérêt de déclaration sociale entre le groupe témoin après et le groupe traité avant l'envoi de courriels.

Afin de compléter l'analyse de la robustesse des effets mesurés, nous prenons en compte l'hétérogénéité inobservable au niveau des établissements, ce qui conduit à calculer de nouveaux écarts-types. Cette hétérogénéité inobservable est notée u_i et nous la considérons comme un effet aléatoire. Le modèle suivant est ainsi estimé par la méthode des moindres carrés généralisés :

$$Y_{i_{g_t}} = \sum \alpha_{1_g}.G_g + \sum \alpha_{2_t}.T_t + \sum \alpha_{3_{g_t}}.G_g.T_t + \beta + u_i + \epsilon_{i_{g_t}}$$

VI. Résultats

1. Estimation non paramétrique de l'effet causal

a. Stock d'emplois

Dans un premier temps, nous identifions le comportement du groupe témoin, qui nous sert de cadre de référence. Pour rappel, ce groupe témoin a été destinataire d'un message neutre, sans contenu incitant à modifier son comportement de déclaration sociale.

Tableau 2 : Déclarations sociales du groupe témoin

Variable	Moyenne	Écart-type	Min	Max
Effectif	11,82	23,56	0	558
Déclaration unique d'embauche en CDI	0,43	1,71	0	61
Déclaration unique d'embauche en CDD court	3,45	39,40	0	1875
Déclaration unique d'embauche en CDD long	0,56	2,71	0	103
Assiette déplafonnée	73631,79	175811,80	0	4465198
Assiette CSG	77395,34	192229,30	0	9772023
Salaire moyen par tête	5510,33	4974,90	0	210895,5

Note : Les indicateurs sont mesurés sur les déclarations sociales postérieures à l'envoi des messages, c'est-à-dire du quatrième trimestre 2013 au quatrième trimestre 2014. Le groupe témoin contient 6898 établissements.

Le stock d'effectif trimestriel moyen de ce groupe a augmenté entre le quatrième trimestre 2013 et le quatrième trimestre 2014 en passant de 11,66 salariés à 11,94 salariés. Les établissements de ce groupe ont déclaré en moyenne 0,43 nouvelle embauche en contrat à durée indéterminée, 3,45 nouvelles embauches en contrat à durée déterminée de courte durée et 0,56 nouvelle embauche en contrat à durée déterminée de longue durée. En moyenne, les établissements du groupe témoin ont versé des rémunérations à hauteur de 73 631,76 euros trimestriellement, et les rémunérations ont eu tendance à augmenter sur la période. Enfin, la rémunération des salariés sur un trimestre était en moyenne de 5510,33 euros.

Les résultats des estimations non paramétriques à l'aide d'un test t de Student sont reportés dans l'ensemble des tableaux 3. Chaque tableau répertorie pour chaque trimestre les différences observées sur les indicateurs pour chaque message envoyé.

Tableau 3 : Tests de comparaison de moyenne sur les indicateurs de stock

Tableau 3-1 : Effectif

Groupe	Trimestre 4 2013	Trimestre 1 2014	Trimestre 2 2014	Trimestre 3 2014	Trimestre 4 2014
Témoin	11,66*** (0,282)	11,61*** (0,283)	11,96*** (0,293)	11,96*** (0,295)	11,94*** (0,298)
Risque de détection	0,383 (0,397)	0,304 (0,398)	0,268 (0,412)	0,377 (0,416)	0,463 (0,421)
Sanctions monétaires	0,563 (0,397)	0,442 (0,398)	0,432 (0,412)	0,628 (0,416)	0,702* (0,420)
Sanctions non monétaires	0,214 (0,397)	0,0926 (0,398)	0,0336 (0,412)	0,103 (0,415)	0,156 (0,420)
Coûts cachés	0,227 (0,396)	0,175 (0,397)	0,135 (0,411)	0,188 (0,414)	0,185 (0,419)
Coûts différés	0,341 (0,395)	0,251 (0,396)	0,202 (0,410)	0,329 (0,413)	0,383 (0,418)
Contribution production bien public	0,405 (0,399)	0,375 (0,400)	0,276 (0,413)	0,339 (0,417)	0,488 (0,422)
« Frauder c'est voler »	0,240 (0,397)	0,160 (0,398)	0,0369 (0,412)	0,160 (0,416)	0,279 (0,420)
Publicité des noms des fraudeurs	-0,0333 (0,397)	-0,128 (0,398)	-0,196 (0,412)	-0,152 (0,415)	-0,118 (0,420)
Mise en danger d'autrui	0,0860 (0,397)	0,0322 (0,398)	0,0364 (0,411)	0,0598 (0,415)	0,104 (0,419)
Norme sociale	**0,823** (0,397)**	**0,835** (0,398)**	**0,846** (0,412)**	**0,953** (0,415)**	**1,010** (0,420)**
Nombre d'observations	76069	75966	74603	73192	71699

Note : La moyenne de l'indicateur est inscrite pour le groupe témoin pour chaque trimestre. Les écarts à ces valeurs sont inscrits pour les groupes traités. Les erreurs-types sont entre parenthèses. *** : p-value inférieure à 0,01 ; ** : p-value inférieure à 0,05 ; * : p-value inférieure à 0,10.

Tableau 3-2 : Déclaration unique d'embauche en contrat à durée indéterminée

Groupe	Trimestre 4 2013	Trimestre 1 2014	Trimestre 2 2014	Trimestre 3 2014	Trimestre 4 2014
Témoin	0,401*** (0,0347)	0,450*** (0,0334)	0,404*** (0,0371)	0,477*** (0,0470)	0,395*** (0,0396)
Risque de détection	0,0217 (0,0489)	-0,00241 (0,0471)	0,0176 (0,0522)	0,0438 (0,0662)	0,00183 (0,0558)
Sanctions monétaires	0,105** (0,0488)	0,0727 (0,0471)	0,0725 (0,0522)	0,0601 (0,0662)	0,0946* (0,0558)
Sanctions non monétaires	0,0783 (0,0488)	0,0402 (0,0471)	0,0224 (0,0522)	0,0688 (0,0661)	0,0896 (0,0557)
Coûts cachés	0,0447 (0,0487)	0,0258 (0,0469)	0,0333 (0,0521)	0,0411 (0,0659)	0,00269 (0,0555)
Coûts différés	**0,111** (0,0486)**	**0,103** (0,0468)**	**0,103** (0,0519)**	**0,144** (0,0658)**	0,0759 (0,0554)
Contribution production bien public	**0,103** (0,0490)**	**0,107** (0,0473)**	0,0725 (0,0524)	0,0699 (0,0664)	0,0488 (0,0559)
« Frauder c'est voler »	0,0722 (0,0488)	0,0106 (0,0471)	0,0238 (0,0522)	0,0460 (0,0662)	0,0101 (0,0558)
Publicité des noms des fraudeurs	0,0355 (0,0489)	0,0272 (0,0471)	0,00272 (0,0522)	0,0594 (0,0662)	0,0208 (0,0557)
Mise en danger d'autrui	0,0514 (0,0488)	0,0232 (0,0470)	0,0342 (0,0522)	0,0134 (0,0661)	-0,0395 (0,0556)
Norme sociale	**0,116** (0,0488)**	0,0533 (0,0470)	0,0813 (0,0522)	0,0666 (0,0661)	0,0286 (0,0557)
Nombre d'observations	76069	75966	74603	73192	71699

Note : La moyenne de l'indicateur est inscrite pour le groupe témoin pour chaque trimestre. Les écarts à ces valeurs sont inscrits pour les groupes traités. Les erreurs-types sont entre parenthèses. *** : p-value inférieure à 0,01 ; ** : p-value inférieure à 0,05 ; * : p-value inférieure à 0,10.

Tableau 3-3 : Déclaration unique d'embauche en contrat à durée déterminée court

Groupe	Trimestre 4 2013	Trimestre 1 2014	Trimestre 2 2014	Trimestre 3 2014	Trimestre 4 2014
Témoin	3,386***	2,876***	3,468***	3,851***	3,717***
	(0,490)	(0,491)	(0,628)	(0,450)	(0,470)
Risque de détection	-0,0691	0,963	1,240	-0,479	0,160
	(0,689)	(0,692)	(0,883)	(0,634)	(0,662)
Sanctions monétaires	-0,426	-0,205	-0,149	-0,587	-0,668
	(0,688)	(0,691)	(0,883)	(0,634)	(0,662)
Sanctions non monétaires	-0,0983	-0,311	-0,316	-0,607	-0,806
	(0,688)	(0,691)	(0,883)	(0,633)	(0,662)
Coûts cachés	0,470	0,570	0,716	0,254	-0,0958
	(0,687)	(0,689)	(0,880)	(0,631)	(0,660)
Coûts différés	-0,789	-0,442	-0,494	-0,692	-0,633
	(0,685)	(0,688)	(0,878)	(0,630)	(0,658)
Contribution production bien public	0,198	0,298	0,231	-0,273	-0,0841
	(0,691)	(0,694)	(0,886)	(0,636)	(0,665)
« Frauder c'est voler »	-1,033	-0,695	-0,855	**-1,091***	**-1,259***
	(0,688)	(0,691)	(0,883)	**(0,634)**	**(0,662)**
Publicité des noms des fraudeurs	0,0790	0,125	0,0762	-0,202	-0,232
	(0,689)	(0,691)	(0,883)	(0,633)	(0,661)
Mise en danger d'autrui	-1,132	-0,781	-0,981	**-1,388****	**-1,416****
	(0,688)	(0,691)	(0,882)	**(0,633)**	**(0,661)**
Norme sociale	-0,201	-0,0701	-0,604	**-1,098***	**-1,118***
	(0,688)	(0,691)	(0,882)	**(0,633)**	**(0,661)**
Nombre d'observations	76069	75966	74603	73192	71699

Note : La moyenne de l'indicateur est inscrite pour le groupe témoin pour chaque trimestre. Les écarts à ces valeurs sont inscrits pour les groupes traités. Les erreurs-types sont entre parenthèses. *** : p-value inférieure à 0,01 ; ** : p-value inférieure à 0,05 ; * : p-value inférieure à 0,10.

Tableau 3-4 : Déclaration unique d'embauche en contrat à durée déterminée long

Groupe	Trimestre 4 2013	Trimestre 1 2014	Trimestre 2 2014	Trimestre 3 2014	Trimestre 4 2014
Témoin	0,440***	0,520***	0,590***	0,819***	0,406***
	(0,0376)	(0,0412)	(0,0468)	(0,0505)	(0,0407)
Risque de détection	0,0470	0,00132	0,0404	0,000886	0,0883
	(0,0530)	(0,0580)	(0,0659)	(0,0711)	(0,0574)
Sanctions monétaires	0,0190	-0,0404	0,0180	0,0783	0,0481
	(0,0529)	(0,0580)	(0,0658)	(0,0711)	(0,0573)
Sanctions non monétaires	-0,00311	-0,0948	0,00635	-0,00360	0,0168
	(0,0529)	(0,0580)	(0,0658)	(0,0710)	(0,0573)
Coûts cachés	0,0240	-0,0636	0,0257	-0,0109	0,0287
	(0,0528)	(0,0578)	(0,0656)	(0,0708)	(0,0571)
Coûts différés	0,0620	-0,000313	0,0658	0,0977	0,0776
	(0,0526)	(0,0577)	(0,0655)	(0,0707)	(0,0570)
Contribution production bien public	0,0793	-0,0227	0,0246	0,0584	0,0905
	(0,0531)	(0,0582)	(0,0661)	(0,0714)	(0,0575)
« Frauder c'est voler »	-0,0140	-0,0459	0,0101	0,00825	-0,0107
	(0,0529)	(0,0580)	(0,0658)	(0,0711)	(0,0573)
Publicité des noms des fraudeurs	0,0492	-0,0353	0,0399	-0,0169	0,0401
	(0,0529)	(0,0580)	(0,0658)	(0,0711)	(0,0572)
Mise en danger d'autrui	-0,00342	0,000120	0,0898	0,0714	0,0448
	(0,0529)	(0,0579)	(0,0658)	(0,0709)	(0,0572)
Norme sociale	**0,0903***	0,00122	0,0339	0,0390	**0,109***
	(0,0529)	(0,0579)	(0,0658)	(0,0710)	**(0,0573)**
Nombre d'observations	76069	75966	74603	73192	71699

Note : La moyenne de l'indicateur est inscrite pour le groupe témoin pour chaque trimestre. Les écarts à ces valeurs sont inscrits pour les groupes traités. Les erreurs-types sont entre parenthèses. *** : p-value inférieure à 0,01 ; ** : p-value inférieure à 0,05 ; * : p-value inférieure à 0,10.

Tableau 3-5 : Assiette déplafonnée

Groupe	Trimestre 4 2013	Trimestre 1 2014	Trimestre 2 2014	Trimestre 3 2014	Trimestre 4 2014
Témoin	75095,4*** (3008,4)	69938,8*** (2462,3)	73220,5*** (2372,5)	71882,6*** (5765,5)	78206,5*** (3056,7)
Risque de détection	1306,8 (4233,5)	1747,6 (3467,4)	1371,8 (3338,7)	1018,6 (8119,7)	1765,7 (4306,3)
Sanctions monétaires	2027,6 (4227,9)	5213,6 (3464,5)	1715,8 (3337,0)	1773,4 (8117,3)	2603,1 (4304,7)
Sanctions non monétaires	1888,1 (4228,8)	-765,2 (3464,4)	-26,85 (3337,4)	4163,2 (8109,4)	1491,1 (4301,4)
Coûts cachés	5333,4 (4218,4)	3512,2 (3454,2)	3568,7 (3326,3)	9917,6 (8084,4)	5472,5 (4287,8)
Coûts différés	93,27 (4207,8)	390,6 (3447,3)	63,20 (3318,3)	-187,7 (8070,4)	363,6 (4278,2)
Contribution production bien public	1298,0 (4247,0)	679,5 (3479,7)	697,7 (3350,2)	81,98 (8144,9)	1368,6 (4319,8)
« Frauder c'est voler »	-277,0 (4229,4)	-361,9 (3464,2)	-1677,3 (3335,8)	-1584,9 (8113,1)	-532,1 (4305,0)
Publicité des noms des fraudeurs	-1576,0 (4232,0)	-1173,0 (3465,0)	-2414,9 (3336,1)	-2708,0 (8110,0)	-2274,1 (4298,1)
Mise en danger d'autrui	**7816,9* (4228,1)**	2624,4 (3462,5)	3409,1 (3333,5)	**14024,9* (8097,8)**	5262,4 (4294,4)
Norme sociale	4759,3 (4226,9)	4262,5 (3463,1)	4529,9 (3335,6)	4017,0 (8106,1)	5729,5 (4299,8)
Nombre d'observations	76069	75966	74603	73192	71699

Note : La moyenne de l'indicateur est inscrite pour le groupe témoin pour chaque trimestre. Les écarts à ces valeurs sont inscrits pour les groupes traités. Les erreurs-types sont entre parenthèses. *** : p-value inférieure à 0,01 ; ** : p-value inférieure à 0,05 ; * : p-value inférieure à 0,10.

Tableau 3-6 : Assiette contribution sociale généralisée

Groupe	Trimestre 4 2013	Trimestre 1 2014	Trimestre 2 2014	Trimestre 3 2014	Trimestre 4 2014
Témoin	78785,0*** (3103,2)	73403,6*** (2555,6)	77720,8*** (2563,9)	75000,2*** (5856,6)	82258,7*** (3166,3)
Risque de détection	1223,3 (4366,9)	1172,5 (3598,8)	1577,4 (3608,1)	592,9 (8248,0)	1539,0 (4460,7)
Sanctions monétaires	1884,5 (4361,1)	4468,9 (3595,8)	2224,5 (3606,2)	2125,3 (8245,6)	2399,7 (4459,0)
Sanctions non monétaires	1924,8 (4362,0)	-1450,6 (3595,7)	23,53 (3606,6)	4201,6 (8237,6)	1255,1 (4455,6)
Coûts cachés	4908,0 (4351,2)	3290,2 (3585,1)	3573,1 (3594,6)	9574,8 (8212,2)	5090,2 (4441,6)
Coûts différés	-46,73 (4340,3)	-33,36 (3578,0)	391,2 (3586,0)	232,7 (8197,9)	260,6 (4431,5)
Contribution production bien public	478,3 (4380,8)	32,69 (3611,7)	260,6 (3620,5)	24,78 (8273,6)	936,0 (4474,7)
« Frauder c'est voler »	-356,5 (4362,7)	-793,4 (3595,6)	-1565,8 (3604,9)	-1654,7 (8241,3)	-705,3 (4459,3)
Publicité des noms des fraudeurs	-1764,4 (4365,3)	-1778,0 (3596,4)	-3145,6 (3605,2)	-3215,4 (8238,2)	-2390,9 (4452,2)
Mise en danger d'autrui	**7575,7* (4361,3)**	1928,0 (3593,8)	3161,1 (3602,4)	**14193,8* (8225,7)**	4955,6 (4448,4)
Norme sociale	4284,4 (4360,0)	3565,5 (3594,4)	4492,7 (3604,6)	4158,0 (8234,2)	5179,5 (4453,9)
Nombre d'observations	76069	75966	74603	73192	71699

Note : La moyenne de l'indicateur est inscrite pour le groupe témoin pour chaque trimestre. Les écarts à ces valeurs sont inscrits pour les groupes traités. Les erreurs-types sont entre parenthèses. *** : p-value inférieure à 0,01 ; ** : p-value inférieure à 0,05 ; * : p-value inférieure à 0,10.

Nous observons que quatre messages sur dix n'ont eu aucun impact significatif sur la décision de déclarer les informations sociales. Il s'agit des messages dont le contenu informait sur les modalités de contrôle de la fraude sociale, les sanctions non financières, l'existence de coûts cachés liés à la fraude et sur la possibilité pour les autorités de rendre public le nom des entreprises recourant au travail illégal et à de mauvaises déclarations de cotisation sociale. Ces messages n'ont pas été perçus comme suffisamment incitatifs pour modifier les comportements.

En ce qui concerne les autres messages, nous constatons que les courriels qui faisaient appel aux interactions sociales et à la moralité contributive ont provoqué davantage de modifications comportementales que les courriels plus coercitifs et qui faisaient appel aux paramètres canoniques des modèles du consentement au prélèvement obligatoire.

Pour rappel, les courriels ont été envoyés fin décembre 2013 et les entreprises devaient déclarer leurs informations sociales pour le dernier trimestre 2013 dans le courant du premier trimestre 2014. L'impact des messages a été plus important lors du trimestre suivant l'envoi de déclaration. Les messages ont donc un impact immédiat sur le comportement de déclaration, mais pour certains d'entre eux cet effet n'est pas durable.

L'indicateur de déclaration sociale qui a subi le plus de modifications est la déclaration d'embauche en CDI. Notons que parmi l'ensemble des effets, seule la déclaration unique d'embauche en CDD de court terme est réduite après l'envoi des courriels, les autres effets étant positifs. Il existerait donc un recours plus important au CDD de court terme pour les établissements par rapport au groupe témoin qui a été réduit à la suite de l'envoi des courriels.

Le courriel décrivant les sanctions financières a augmenté la déclaration en CDI au cours des derniers trimestres 2013 et 2014. Les entreprises de ce groupe ont également augmenté la déclaration de leur effectif au cours du quatrième trimestre 2014.

Le message qui apporte des précisions sur les coûts différés engendrés par la fraude, comme l'interdiction d'accès à des marchés publics, augmente la déclaration d'embauche en CDI, du quatrième trimestre 2013 au troisième trimestre 2014. C'est donc un effet durable, et il semble que ce contenu soit un levier spécifique pour une meilleure déclaration des employés en CDI.

Le courriel qui détaille l'utilité de la contribution sociale, en précisant les destinations du financement des cotisations sociales, augmente également la déclaration d'embauche en CDI du quatrième trimestre 2013 au premier trimestre 2014. L'ordre de grandeur est comparable avec le message précédent.

Quant à celui qui développe l'idée selon laquelle « la fraude est du vol », ce courriel agit négativement sur la déclaration d'embauche en CDD de court terme au cours du troisième et du quatrième trimestre 2014.

Enfin, le message qui insiste sur la notion de norme sociale du consentement à la contribution sociale est celui qui produit le plus d'impact sur la déclaration sociale. Au cours du quatrième trimestre 2013, il provoque une augmentation de la déclaration d'embauche en CDI et en CDD de long terme, ainsi que de la déclaration de l'effectif. Cet effet se maintient pendant toute la période d'observation post-traitement. Le message diminue la déclaration d'embauche en CDD de court terme au cours des deux derniers trimestres de l'année 2014. Ce message augmente la déclaration d'embauche en CDD de long terme au cours du dernier trimestre 2014.

Pour synthétiser les résultats de ces estimations non paramétriques de l'impact des courriels sur les indicateurs de stock, nous inscrivons les signes des effets dans le tableau 3bis. L'estimation paramétrique des effets permettra de confirmer ou d'infirmer ces constats.

Tableau 3bis : Synthèse des effets

Indicateur / Groupe		Risque de détection	Sanctions monétaires	Sanctions non monétaires	Coûts cachés	Coûts différés	Contribution à la production d'un bien public	« Frauder c'est voler »	Publicité du nom des fraudeurs	Mise en danger d'autrui	Norme sociale
T42013	EFF										+
	DUECDI		+			+	+				+
	DUECDDC									-	
	DUECDDL										+
	ASSDEP									+	
	ASSCSG									+	
T12014	EFF										+
	DUECDI					+	+				
	DUECDDC										
	DUECDDL										
	ASSDEP										
	ASSCSG										
T22014	EFF										+
	DUECDI					+					
	DUECDDC										
	DUECDDL										
	ASSDEP										
	ASSCSG										
T32014	EFF										+
	DUECDI					+					
	DUECDDC							-		-	-
	DUECDDL										
	ASSDEP									+	
	ASSCSG									+	
T42014	EFF		+								+
	DUECDI		+								
	DUECDDC							-		-	-
	DUECDDL										+
	ASSDEP										
	ASSCSG										
Total d'effet		0	3	0	0	4	2	2	0	7	10

b. Flux d'emplois

Nous pouvons décrire les flux d'emplois pour le groupe témoin avant de commenter les différences que l'on observe pour chaque groupe d'établissements traité (tableau 4). Nous remarquons que la création d'emploi était supérieure à la destruction d'emploi. L'ensemble des variations positives de stock d'emploi entre 2013 et 2014 rapporté à la taille du groupe est de 15,7 %, ce qui correspond à 1,68 nouvelles entrées pour un stock de 10,73 salariés en moyenne. L'ensemble des variations négatives de stock d'emploi entre 2013 et 2014 rapporté à la taille du groupe est de 14,9 %, ce qui correspond à 1,60 nouvelles sorties pour un stock de 10,73 salariés en moyenne.

Tableau 4 : Flux d'emplois pour le groupe témoin en 2014

Variable	Moyenne	Écart-type	Min	Max
Création d'emploi	1,68	15,04	0	1111
Destruction d'emploi	1,60	6,42	0	189
Taille	10,73	25,88	0	628,5

Note : le groupe témoin contient 6898 observations.

Les résultats issus des tests de nullité de différences entre le groupe témoin et les groupes traités figurent dans le tableau 5. Comparativement aux indicateurs de stock, seuls quatre messages ont modifié le comportement de création et de destruction d'emploi. De plus, tous les effets que nous observons sont des résultats négatifs. Enfin, c'est davantage les flux d'emplois en CDD qui ont été modifiés à la suite de l'envoi des messages. Dans cette analyse, il n'est pas possible de distinguer la durée des CDD.

Le message qui décrit l'architecture des sanctions monétaires réduit significativement le taux de création d'emploi en CDD.

Le courriel qui insiste sur l'existence de coûts cachés liés à la fraude sociale influence également négativement le taux de création en CDD et négativement le taux de croissance du nombre d'employés en CDD. Cela veut dire qu'il y a eu plus de destructions que de créations d'emploi en CDD et que cela a conduit a diminué l'emploi en CDD.

La réduction de la croissance des emplois en CDD s'observe également pour les établissements ayant reçu le message contenant une description des sanctions non financières, c'est-à-dire les peines d'emprisonnement et ceux qui ont été destinataires du message qui traite du respect de la norme sociale au prélèvement obligatoire réduit.

Tableau 5 : Tests de comparaison de moyenne sur les indicateurs de flux d'emplois en 2014

Groupe	Taux de création	*dont CDI*	*dont CDD*	Taux de destruction	*dont CDI*	*dont CDD*	Taux de croissance	*dont CDI*	*dont CDD*
Témoin	15,65***	9,195***	6,457***	14,89***	9,873***	5,020***	0,758	-0,678	1,437*
	(0,983)	(0,659)	(0,630)	(0,835)	(0,616)	(0,476)	(1,262)	(0,916)	(0,795)
Risque de détection	**-2,365***	**-1,731***	-0,634	-0,404	-0,526	0,122	-1,961	-1,205	-0,756
	(1,383)	**(0,928)**	(0,887)	(1,175)	(0,868)	(0,669)	(1,776)	(1,289)	(1,119)
Sanctions monétaires	-2,157	-0,628	**-1,529***	-0,535	-0,436	-0,0990	-1,623	-0,193	-1,430
	(1,382)	(0,927)	**(0,887)**	(1,175)	(0,867)	(0,669)	(1,774)	(1,288)	(1,118)
Sanctions non monétaires	-2,152	-0,733	-1,420	1,301	0,270	1,030	**-3,453***	-1,003	**-2,450****
	(1,382)	(0,927)	(0,887)	(1,175)	(0,867)	(0,669)	**(1,775)**	(1,289)	**(1,118)**
Coûts cachés	-**2,300***	-0,386	-**1,914****	0,387	-0,0357	0,422	-2,687	-0,350	**-2,336****
	(1,378)	(0,925)	**(0,884)**	(1,172)	(0,865)	(0,667)	(1,770)	(1,285)	**(1,115)**
Coûts différés	-1,314	-0,182	-1,133	-0,357	-0,613	0,256	-0,957	0,432	-1,389
	(1,376)	(0,923)	(0,883)	(1,169)	(0,863)	(0,666)	(1,766)	(1,283)	(1,113)
Contribution production bien public	-1,418	-0,499	-0,919	0,814	0,376	0,437	-2,232	-0,875	-1,357
	(1,387)	(0,931)	(0,890)	(1,179)	(0,870)	(0,671)	(1,781)	(1,293)	(1,122)
« Frauder c'est voler »	-0,345	1,022	-1,367	-0,376	-0,214	-0,163	0,0315	1,235	-1,204
	(1,383)	(0,928)	(0,888)	(1,176)	(0,868)	(0,669)	(1,776)	(1,290)	(1,119)
Publicité des noms des fraudeurs	-1,518	-0,567	-0,951	-0,281	0,000472	-0,282	-1,237	-0,567	-0,669
	(1,383)	(0,927)	(0,887)	(1,175)	(0,867)	(0,669)	(1,775)	(1,289)	(1,118)
Mise en danger d'autrui	-0,795	0,00621	-0,801	0,863	0,555	0,308	-1,658	-0,549	-1,109
	(1,381)	(0,926)	(0,886)	(1,174)	(0,866)	(0,668)	(1,773)	(1,287)	(1,117)
Norme sociale	**-2,467***	-1,085	-1,381	0,340	-0,188	0,528	-2,807	-0,897	**-1,910***
	(1,382)	(0,927)	(0,887)	(1,174)	(0,867)	(0,669)	(1,774)	(1,288)	**(1,118)**
Nombre d'observations	77476	77476	77476	77476	77476	77476	77476	77476	77476

Note : La moyenne de l'indicateur est inscrite pour le groupe témoin pour l'année 2014. Les écarts à ces valeurs sont inscrits pour les groupes traités. Les erreurs-types sont entre parenthèses. *** : p-value inférieure à 0,01 ; ** : p-value inférieure à 0,05 ; * : p-value inférieure à 0,10.

2. Estimation paramétrique de l'effet causal par double différence

Après avoir identifié ces nombreuses différences significatives qui correspondent à l'estimation « naïve » de l'intention de traiter, il est judicieux de vérifier si ce que nous observons est l'effet causal du traitement ou bien une erreur de mesure. Un moyen de répondre à cette interrogation est d'utiliser les informations antérieures à l'envoi des messages. Ainsi, en tenant compte de la dimension temporelle, il est possible d'identifier des différences entre les groupes avant le traitement afin de corriger les différences constatées après le traitement.

Nous présentons les résultats issus des estimations qui tiennent compte de l'hétérogénéité inobservable. Comme nous l'avons précisé dans la partie stratégie d'identification, nous considérons cette hétérogénéité inobservable comme étant aléatoire.

Nous produisons une estimation de l'effet causal du traitement grâce à une double différence entre d'une part la différence postérieure au traitement et d'autre part la différence antérieure au traitement entre le groupe témoin et chaque groupe traité. Autrement dit, nous estimons chaque double différence en tenant compte de chaque trimestre selon la décomposition suivante :

$$\left(\overline{Y_{\iota test_{après}}} - \overline{Y_{\iota témoin_{après}}}\right) - \left(\overline{Y_{\iota test_{avant}}} - \overline{Y_{\iota témoin_{avant}}}\right)$$

où l'indice *après* correspond à tous les trimestres (année) postérieurs à l'envoi des messages pour les indicateurs de stock (flux) et l'indice *avant* correspond aux trois trimestres (à l'année) antérieurs à l'expérimentation pour les indicateurs de stock (flux).

a. Stock d'emploi

L'estimation de l'effet « pur » des courriels qui tient compte des comportements de déclarations sociales en stock antérieurs à l'expérimentation est réalisée en utilisant l'information des trois premiers trimestres de l'année 2013 avant l'expérimentation (voir tableau 6).

Tableau 6 : Estimation de l'effet de l'intention de traiter par différence de différences sur les indicateurs de stock

Groupe		Effectif	Déclaration d'embauche en CDI	Déclaration d'embauche en CDD court	Déclaration d'embauche en CDD long	Assiette déplafonnée	Assiette CSG
Témoin		11,72***	0,517***	3,162***	0,649***	69968,2***	73727,6***
		(0,165)	(0,0229)	(0,279)	(0,0251)	(2115,0)	(2174,8)
Trimestre 4 2013		-0,0643	**-0,116****	0,224	**-0,209*****	5127,2	5057,4
		(0,330)	**(0,0457)**	(0,557)	**(0,0502)**	(4221,4)	(4340,8)
Trimestre 1 2014		-0,112	-0,0676	-0,287	**-0,129****	-29,34	-324,0
		(0,329)	(0,0457)	(0,556)	**(0,0501)**	(4218,8)	(4338,2)
Trimestre 2 2014		0,239	**-0,114****	0,306	-0,0590	3252,3	3993,3
		(0,332)	**(0,0460)**	(0,561)	(0,0505)	(4251,3)	(4371,6)
Trimestre 3 2014		0,236	-0,0400	0,689	**0,171*****	1914,4	1272,7
		(0,334)	(0,0463)	(0,564)	**(0,0509)**	(4279,7)	(4400,8)
Trimestre 4 2014		0,216	**-0,122*****	0,555	**-0,243*****	**8238,3***	**8531,1***
		(0,337)	**(0,0467)**	(0,568)	**(0,0512)**	**(4311,5)**	**(4433,5)**
Risque de détection		0,373	-0,0199	-0,251	0,0116	1933,6	1621,3
		(0,232)	(0,0322)	(0,393)	(0,0354)	(2977,1)	(3061,3)
Sanctions monétaires		**0,594****	0,0501	-0,434	0,00334	2995,2	2908,3
		(0,232)	(0,0322)	(0,392)	(0,0353)	(2974,2)	(3058,4)
Sanctions non monétaires		0,272	0,0187	-0,341	-0,00841	1552,1	1269,6
		(0,232)	(0,0322)	(0,392)	(0,0354)	(2975,2)	(3059,4)
Coûts cachés		0,235	0,0169	0,861**	-0,00346	**6145,5****	**6192,7****
		(0,232)	(0,0321)	(0,391)	(0,0353)	**(2966,6)**	**(3050,5)**
Coûts différés		0,299	**0,0577***	**-0,660***	0,0420	-37,84	-250,2
		(0,231)	**(0,0321)**	**(0,390)**	(0,0352)	(2961,4)	(3045,2)
Contribution production bien public		0,375	0,0528	0,243	0,00496	560,9	-76,74
		(0,233)	(0,0323)	(0,394)	(0,0355)	(2987,0)	(3071,5)
« Frauder c'est voler »		0,233	-0,00614	**-0,932****	0,000336	-687,1	-990,2
		(0,232)	(0,0322)	**(0,393)**	(0,0354)	(2977,0)	(3061,2)
Publicité des noms des fraudeurs		-0,0318	-0,00537	0,182	0,0102	-1459,2	-2209,1
		(0,233)	(0,0322)	(0,393)	(0,0354)	(2978,0)	(3062,3)
Mise en danger d'autrui		0,0784	-0,00307	**-0,822****	0,0222	**8255,8*****	**8062,2*****
		(0,232)	(0,0322)	**(0,392)**	(0,0353)	**(2973,3)**	**(3057,4)**
Norme sociale		**0,771*****	**0,0595***	-0,370	0,0141	3778,7	3488,8
		(0,232)	**(0,0322)**	(0,392)	(0,0353)	(2972,4)	(3056,4)
Risque de détection	Trimestre 4 2013	0,0103	0,0416	0,182	0,0355	-626,8	-398,0
		(0,464)	(0,0643)	(0,783)	(0,0706)	(5940,9)	(6108,9)
	Trimestre 1 2014	-0,0682	0,0175	1,214	-0,0102	-186,0	-448,8
		(0,464)	(0,0643)	(0,783)	(0,0706)	(5940,3)	(6108,4)
	Trimestre 2 2014	-0,105	0,0375	**1,491***	0,0289	-561,7	-43,83
		(0,467)	(0,0648)	**(0,789)**	(0,0711)	(5983,0)	(6152,3)
	Trimestre 3 2014	0,00471	0,0638	-0,228	-0,0107	-914,9	-1028,3
		(0,471)	(0,0652)	(0,795)	(0,0716)	(6026,5)	(6197,0)
	Trimestre 4 2014	0,0906	0,0217	0,411	0,0767	-167,9	-82,25
		(0,474)	(0,0657)	(0,801)	(0,0722)	(6072,8)	(6244,6)
Sanctions monétaires	Trimestre 4 2013	-0,0303	0,0547	0,00791	0,0157	-967,7	-1023,9
		(0,463)	(0,0642)	(0,782)	(0,0705)	(5933,5)	(6101,4)
	Trimestre 1 2014	-0,152	0,0225	0,230	-0,0437	2218,4	1560,6
		(0,463)	(0,0642)	(0,783)	(0,0705)	(5935,2)	(6103,1)
	Trimestre 2 2014	-0,162	0,0224	0,285	0,0147	-1279,5	-683,9
		(0,467)	(0,0647)	(0,788)	(0,0711)	(5979,3)	(6148,5)
	Trimestre 3 2014	0,0339	0,00997	-0,153	0,0749	-1221,8	-783,1
		(0,470)	(0,0652)	(0,794)	(0,0716)	(6023,7)	(6194,1)
	Trimestre 4 2014	0,108	0,0445	-0,234	0,0448	-392,1	-508,6
		(0,474)	(0,0657)	(0,800)	(0,0721)	(6069,6)	(6241,3)
Sanctions non monétaires	Trimestre 4 2013	-0,0578	0,0595	0,243	0,00530	336,0	655,1
		(0,463)	(0,0642)	(0,783)	(0,0705)	(5935,0)	(6102,9)
	Trimestre 1 2014	-0,180	0,0214	0,0300	-0,0864	-2317,3	-2720,2
		(0,463)	(0,0642)	(0,783)	(0,0705)	(5935,5)	(6103,5)
	Trimestre 2 2014	-0,239	0,00368	0,0248	0,0148	-1579,0	-1246,1
		(0,467)	(0,0647)	(0,789)	(0,0711)	(5980,3)	(6149,5)
	Trimestre 3 2014	-0,169	0,0501	-0,265	0,00481	2611,1	2932,0
		(0,470)	(0,0652)	(0,794)	(0,0715)	(6019,8)	(6190,1)

Groupe		Effectif	Déclaration d'embauche en CDI	Déclaration d'embauche en CDD court	Déclaration d'embauche en CDD long	Assiette déplafonnée	Assiette CSG
	Trimestre 4 2014	-0,116 (0,474)	0,0708 (0,0657)	-0,464 (0,800)	0,0252 (0,0721)	-60,98 (6066,6)	-14,58 (6238,3)
Coûts cachés	Trimestre 4 2013	-0,00795 (0,462)	0,0279 (0,0641)	-0,390 (0,781)	0,0274 (0,0704)	-812,1 (5919,7)	-1284,7 (6087,1)
	Trimestre 1 2014	-0,0599 (0,462)	0,00893 (0,0641)	-0,291 (0,780)	-0,0601 (0,0703)	-2633,4 (5918,1)	-2902,5 (6085,6)
	Trimestre 2 2014	-0,0995 (0,465)	0,0165 (0,0645)	-0,145 (0,786)	0,0292 (0,0708)	-2576,8 (5961,1)	-2619,6 (6129,7)
	Trimestre 3 2014	-0,0469 (0,469)	0,0242 (0,0650)	-0,607 (0,791)	-0,00748 (0,0713)	3772,0 (6001,5)	3382,1 (6171,3)
	Trimestre 4 2014	-0,0501 (0,472)	-0,0142 (0,0655)	-0,956 (0,797)	0,0322 (0,0719)	-673,0 (6047,8)	-1102,4 (6219,0)
Coûts différés	Trimestre 4 2013	0,0414 (0,461)	0,0533 (0,0639)	-0,129 (0,779)	0,0200 (0,0702)	131,1 (5906,0)	203,4 (6073,1)
	Trimestre 1 2014	-0,0481 (0,461)	0,0453 (0,0639)	0,219 (0,779)	-0,0423 (0,0702)	428,5 (5906,7)	216,8 (6073,8)
	Trimestre 2 2014	-0,0976 (0,464)	0,0457 (0,0644)	0,167 (0,784)	0,0238 (0,0707)	101,0 (5947,7)	641,3 (6116,0)
	Trimestre 3 2014	0,0300 (0,468)	0,0868 (0,0648)	-0,0320 (0,790)	0,0557 (0,0712)	-149,9 (5991,1)	482,8 (6160,6)
	Trimestre 4 2014	0,0840 (0,471)	0,0182 (0,0653)	0,0277 (0,796)	0,0355 (0,0717)	401,4 (6035,0)	510,8 (6205,7)
Contribution production bien public	Trimestre 4 2013	0,0302 (0,465)	0,0501 (0,0645)	-0,0447 (0,786)	0,0744 (0,0708)	737,2 (5960,0)	555,1 (6128,7)
	Trimestre 1 2014	0,0000409 (0,465)	0,0545 (0,0645)	0,0557 (0,786)	-0,0276 (0,0708)	118,6 (5961,2)	109,4 (6129,8)
	Trimestre 2 2014	-0,0989 (0,469)	0,0197 (0,0650)	-0,0116 (0,792)	0,0197 (0,0713)	136,9 (6003,5)	337,3 (6173,3)
	Trimestre 3 2014	-0,0361 (0,472)	0,0171 (0,0654)	-0,516 (0,797)	0,0534 (0,0718)	-478,9 (6045,5)	101,5 (6216,6)
	Trimestre 4 2014	0,113 (0,476)	-0,00400 (0,0659)	-0,327 (0,803)	0,0855 (0,0724)	807,8 (6092,1)	1012,7 (6264,5)
« Frauder c'est voler »	Trimestre 4 2013	0,00701 (0,464)	0,0783 (0,0642)	-0,100 (0,783)	-0,0143 (0,0706)	410,1 (5936,5)	633,7 (6104,4)
	Trimestre 1 2014	-0,0724 (0,464)	0,0167 (0,0642)	0,237 (0,783)	-0,0462 (0,0705)	325,2 (5936,2)	196,9 (6104,2)
	Trimestre 2 2014	-0,196 (0,467)	0,0299 (0,0647)	0,0774 (0,788)	0,00974 (0,0711)	-990,2 (5979,0)	-575,6 (6148,2)
	Trimestre 3 2014	-0,0733 (0,470)	0,0522 (0,0652)	-0,159 (0,794)	0,00791 (0,0716)	-897,9 (6022,7)	-664,4 (6193,1)
	Trimestre 4 2014	0,0466 (0,474)	0,0163 (0,0657)	-0,327 (0,801)	-0,0110 (0,0722)	154,9 (6071,3)	284,9 (6243,1)
Publicité des noms des fraudeurs	Trimestre 4 2013	-0,00152 (0,464)	0,0409 (0,0643)	-0,103 (0,783)	0,0390 (0,0706)	-116,7 (5939,7)	444,8 (6107,8)
	Trimestre 1 2014	-0,0961 (0,464)	0,0326 (0,0643)	-0,0576 (0,783)	-0,0455 (0,0706)	286,3 (5937,7)	431,1 (6105,7)
	Trimestre 2 2014	-0,164 (0,467)	0,00809 (0,0647)	-0,106 (0,788)	0,0297 (0,0711)	-955,7 (5979,9)	-936,5 (6149,1)
	Trimestre 3 2014	-0,120 (0,470)	0,0647 (0,0652)	-0,384 (0,794)	-0,0270 (0,0716)	-1248,8 (6021,5)	-1006,3 (6191,9)
	Trimestre 4 2014	-0,0859 (0,474)	0,0262 (0,0656)	-0,414 (0,800)	0,0300 (0,0721)	-814,9 (6064,5)	-181,8 (6236,1)
Mise en danger d'autrui	Trimestre 4 2013	0,00764 (0,463)	0,0544 (0,0642)	-0,310 (0,782)	-0,0256 (0,0705)	-438,9 (5933,2)	-486,5 (6101,1)
	Trimestre 1 2014	-0,0462 (0,463)	0,0263 (0,0642)	0,0408 (0,782)	-0,0221 (0,0705)	-5631,4 (5932,2)	-6134,2 (6100,0)
	Trimestre 2 2014	-0,0419 (0,466)	0,0373 (0,0647)	-0,159 (0,788)	0,0676 (0,0710)	-4846,8 (5974,1)	-4901,1 (6143,1)
	Trimestre 3 2014	-0,0185 (0,469)	0,0165 (0,0651)	-0,566 (0,793)	0,0492 (0,0715)	5769,1 (6012,3)	6131,6 (6182,4)
	Trimestre 4 2014	0,0254 (0,473)	-0,0364 (0,0656)	-0,594 (0,799)	0,0227 (0,0720)	-2993,4 (6058,2)	-3106,7 (6229,6)

Groupe		Effectif	Déclaration d'embauche en CDI	Déclaration d'embauche en CDD court	Déclaration d'embauche en CDD long	Assiette déplafonnée	Assiette CSG
Norme sociale	Trimestre 4 2013	0,0511 (0,463)	0,0567 (0,0642)	0,169 (0,782)	0,0762 (0,0705)	980,7 (5931,5)	795,6 (6099,3)
	Trimestre 1 2014	0,0632 (0,463)	-0,00614 (0,0642)	0,300 (0,782)	-0,0129 (0,0705)	483,8 (5932,5)	76,72 (6100,4)
	Trimestre 2 2014	0,0741 (0,467)	0,0218 (0,0647)	-0,234 (0,788)	0,0198 (0,0710)	751,2 (5976,4)	1003,9 (6145,5)
	Trimestre 3 2014	0,182 (0,470)	0,00708 (0,0651)	-0,728 (0,793)	0,0248 (0,0715)	238,3 (6016,5)	669,2 (6186,7)
	Trimestre 4 2014	0,238 (0,473)	-0,0308 (0,0656)	-0,748 (0,799)	0,0952 (0,0721)	1950,8 (6063,5)	1690,7 (6235,0)
Nombre d'observations		597984	597984	597984	597984	597984	597984

Note : La moyenne de l'indicateur est inscrite pour le groupe témoin pour l'année 2014. Les écarts à ces valeurs sont inscrits pour les groupes traités. Les erreurs-types sont entre parenthèses. *** : p-value inférieure à 0,01 ; ** : p-value inférieure à 0,05 ; * : p-value inférieure à 0,10.

Dans un premier temps, avant d'identifier des différences de comportement liées à l'envoi des messages qui tiennent compte des différences antérieures à l'expérimentation, il est judicieux d'identifier les tendances des indicateurs de déclarations sociales du groupe témoin. Ainsi, l'effectif trimestriel a augmenté sur la période d'analyse, tandis que les déclarations d'embauche en CDI et CDD de long terme ont diminué. La déclaration d'embauche en CDD de court terme a augmenté. La rémunération, mesurée à partir de l'assiette déplafonnée, ou à partir de l'assiette CSG, a également augmenté.

Ensuite, nous constatons des différences significatives sur les indicateurs d'intérêt avant le début de l'expérimentation, ce qui correspond aux différences inscrites de la septième ligne à la seizième ligne du tableau 6. Ces différences mettent en exergue le fait qu'il existait des comportements de déclarations sociales hétérogènes entre le groupe témoin et les groupes traités avant l'envoi des courriels. Les établissements ayant reçu le message décrivant les sanctions financières déclaraient davantage d'effectifs que le groupe témoin. Les établissements qui ont reçu des informations sur les coûts cachés de la fraude déclaraient plus d'embauches en CDD de court terme et davantage de rémunération. Les établissements ayant reçu le message signalant l'existence de coûts différés après la détection d'une fraude déclaraient davantage d'embauche en CDI, mais moins d'embauches de CDD de court terme. Le groupe d'établissements qui était destinataire du courriel insistant sur la notion de « vol » déclarait moins d'embauches en CDD de court terme. Les établissements qui ont reçu des informations sur les conséquences de la fraude sociale comme la mise en danger d'autrui déclaraient également moins d'embauches en CDD de court terme et rémunérait plus ses salariés que le groupe témoin. Enfin, le groupe traité par un courriel contenant l'idée que la contribution sociale est une norme respectée par la majorité des entreprises avait un effectif supérieur au groupe témoin et déclarait davantage d'embauches en CDI.

Les estimations des doubles différences trimestrielles nuancent ainsi de nombreuses estimations « naïves » des intentions de traiter des messages sur les indicateurs considérés. La mesure de l'effet pur de l'envoi des courriels, purgé des différences antérieures de déclarations sociales ne fait pas apparaître de différences significatives entre le groupe témoin et les groupes traités. Seul le message qui insiste sur l'existence de vérifications renforcées sur les informations sociales produit une hausse de la déclaration d'embauche en CDD de courte durée au deuxième trimestre 2014. C'est l'unique effet que nous sommes en mesure de capturer parmi les 385 tests opérés en double différence. Par conséquent, le fait que nous ayons observé des comportements différents après l'envoi des messages ne peut pas directement être associé à l'envoi des messages.

b. Flux d'emploi

L'estimation de l'effet « pur » des courriels qui tient compte des flux d'emploi antérieurs à l'expérimentation est réalisée en utilisant l'information de l'année 2013, c'est-à-dire une année avant l'expérimentation (voir tableau 7).

Avant de décrire les différences de flux d'emploi entre le groupe témoin et les différents groupes traités, nous analysons les tendances sur les indicateurs de flux d'emploi du groupe témoin, entre 2013 et 2014. Ainsi, l'indicateur de création d'emploi a diminué sur la période, de 0,94 point, notamment dû à une diminution importante et significative de la création d'emploi en CDI, tandis que la création de CDD augmentait. Le taux de destruction a augmenté de 2,5 points entre 2013 et 2014, à cause d'une plus grande destruction de l'emploi en CDI. Ainsi, la croissance de l'emploi a faiblement augmenté grâce à une plus importante augmentation de l'embauche en CDD par rapport à la destruction des emplois en CDD.

Avant l'expérimentation, à l'instar des différences de comportement de déclarations sociales en stock, il existe des différences sur les flux d'emploi. Elles s'observent pour le groupe qui a reçu une description des sanctions non monétaires sur les indicateurs de création d'emploi en CDI, et de croissance de l'emploi en CDI.

Tandis que pour les indicateurs de stock, le comportement antérieur pouvait nuancer l'estimation naïve de l'intention de traiter, pour les indicateurs de flux on observe un effet significatif de l'estimation pure de l'intention de traiter. C'est un résultat intéressant, car dans ce contexte, le fait de mesurer un écart significatif en double différence permet d'inférer que l'envoi des messages a joué pleinement le rôle de modification du comportement.

Le groupe d'établissements ayant reçu le message insistant sur les sanctions non financières a eu un taux de croissance de l'emploi, à la fois en CDI et en CDD après l'envoi des courriels significativement inférieur à celui du groupe témoin. Cette plus faible croissance de l'emploi se traduit à la fois par une réduction significative de la création d'emploi en CDI et en CDD, c'est-à-dire que ce groupe d'établissements a eu recours à moins d'embauches en CDI et CDD comparativement au groupe témoin.

Le groupe ayant reçu un message contenant l'idée que la fraude sociale est du vol a réduit significativement le taux de création d'emploi en CDD.

Le groupe ayant reçu un message incitant à respecter la norme sociale du consentement au prélèvement pour lequel le taux de création en CDD était significativement inférieur à celui du groupe témoin voit également son taux de création d'emploi en CDD être réduit. La création en CDD étant inférieure à la destruction en CDD, son taux de croissance de l'emploi en CDD a davantage diminué par rapport au groupe témoin.

Tableau 7 : Estimation de l'effet de l'intention de traiter par différence de différences sur les indicateurs de flux

Groupe		Taux de création	*dont CDI*	*dont CDD*	Taux de destruction	*dont CDI*	*dont CDD*	Taux de croissance	*dont CDI*	*dont CDD*
Témoin		16,59***	11,50***	5,090***	12,40***	6,543***	5,860***	4,192***	4,962***	-0,770
		(0,987)	(0,727)	(0,552)	(0,877)	(0,589)	(0,585)	(1,297)	(0,948)	(0,810)
Année 2014		-0,943	**-2,310****	**1,367***	**2,491****	**3,331*****	-0,839	**-3,434***	**-5,641*****	**2,207***
		(1,283)	**(0,958)**	**(0,752)**	**(1,141)**	**(0,782)**	(0,797)	**(1,834)**	**(1,341)**	**(1,145)**
Risque de détection		-0,0832	-0,470	0,386	-0,581	-0,409	-0,172	0,498	-0,0605	0,559
		(1,390)	(1,024)	(0,777)	(1,234)	(0,829)	(0,823)	(1,825)	(1,334)	(1,140)
Sanctions monétaires		0,443	0,841	-0,398	0,858	-0,0942	0,953	-0,416	0,935	-1,351
		(1,389)	(1,023)	(0,777)	(1,234)	(0,828)	(0,822)	(1,824)	(1,334)	(1,139)
Sanctions non monétaires		**3,114****	**2,237****	0,877	0,246	-0,239	0,484	2,868	**2,476***	0,392
		(1,389)	**(1,023)**	(0,777)	(1,234)	(0,828)	(0,823)	(1,824)	**(1,334)**	(1,139)
Coûts cachés		-1,270	-0,746	-0,524	0,0317	-0,198	0,230	-1,301	-0,548	-0,754
		(1,385)	(1,020)	(0,775)	(1,230)	(0,826)	(0,820)	(1,819)	(1,330)	(1,136)
Coûts différés		-1,388	-1,286	-0,102	-0,853	-0,308	-0,545	-0,536	-0,978	0,443
		(1,383)	(1,018)	(0,773)	(1,228)	(0,824)	(0,819)	(1,815)	(1,328)	(1,134)
Contribution production bien public		0,381	0,209	0,172	0,269	-0,111	0,379	0,113	0,320	-0,207
		(1,394)	(1,027)	(0,780)	(1,238)	(0,831)	(0,826)	(1,831)	(1,339)	(1,143)
« Frauder c'est voler »		0,469	0,0599	0,409	-0,0981	-0,390	0,292	0,567	0,450	0,117
		(1,390)	(1,024)	(0,777)	(1,235)	(0,829)	(0,823)	(1,825)	(1,335)	(1,140)
Publicité des noms des fraudeurs		1,049	1,117	-0,0681	-0,197	-0,524	0,327	1,246	1,641	-0,395
		(1,389)	(1,024)	(0,777)	(1,234)	(0,828)	(0,823)	(1,824)	(1,334)	(1,139)
Mise en danger d'autrui		1,472	1,101	0,371	-0,124	0,208	-0,332	1,596	0,893	0,703
		(1,388)	(1,022)	(0,776)	(1,233)	(0,827)	(0,822)	(1,822)	(1,332)	(1,138)
Norme sociale		0,620	-0,0466	0,667	-1,007	-0,441	-0,566	1,627	0,394	1,233
		(1,388)	(1,023)	(0,776)	(1,233)	(0,828)	(0,822)	(1,823)	(1,333)	(1,138)
Risque de détection	Année 2014	-2,282	-1,261	-1,021	0,177	-0,117	0,294	-2,459	-1,145	-1,315
		(1,805)	(1,349)	(1,058)	(1,606)	(1,100)	(1,122)	(2,581)	(1,887)	(1,612)
Sanctions monétaires	Année 2014	-2,600	-1,469	-1,131	-1,393	-0,341	-1,052	-1,207	-1,128	-0,0794
		(1,804)	(1,348)	(1,058)	(1,605)	(1,099)	(1,121)	(2,579)	(1,886)	(1,610)
Sanctions non monétaires	Année 2014	**-5,266*****	**-2,970****	**-2,296****	1,055	0,509	0,546	**-6,321****	**-3,479***	**-2,843***
		(1,804)	**(1,348)**	**(1,058)**	(1,605)	(1,100)	(1,121)	**(2,579)**	**(1,886)**	**(1,611)**
Coûts cachés	Année	-1,030	0,360	-1,390	0,355	0,162	0,193	-1,385	0,197	-1,583

Groupe		Taux de création	*dont CDI*	*dont CDD*	Taux de destruction	*dont CDI*	*dont CDD*	Taux de croissance	*dont CDI*	*dont CDD*
	2014	(1,800)	(1,345)	(1,055)	(1,601)	(1,097)	(1,118)	(2,572)	(1,881)	(1,606)
Coûts différés	Année 2014	0,0737 (1,796)	1,105 (1,342)	-1,031 (1,053)	0,495 (1,598)	-0,306 (1,095)	0,801 (1,116)	-0,422 (2,567)	1,410 (1,877)	-1,832 (1,603)
Contribution production bien public	Année 2014	-1,799 (1,811)	-0,708 (1,353)	-1,091 (1,062)	0,545 (1,611)	0,487 (1,104)	0,0580 (1,125)	-2,345 (2,589)	-1,195 (1,893)	-1,149 (1,617)
« Frauder c'est voler »	Année 2014	-0,814 (1,806)	0,962 (1,349)	**-1,776* (1,059)**	-0,278 (1,607)	0,176 (1,101)	-0,454 (1,122)	-0,536 (2,581)	0,786 (1,888)	-1,321 (1,612)
Publicité des noms des fraudeurs	Année 2014	-2,567 (1,805)	-1,684 (1,349)	-0,883 (1,058)	-0,0840 (1,606)	0,525 (1,100)	-0,609 (1,122)	-2,483 (2,580)	-2,209 (1,887)	-0,274 (1,611)
Mise en danger d'autrui	Année 2014	-2,267 (1,803)	-1,094 (1,347)	-1,172 (1,057)	0,987 (1,604)	0,347 (1,099)	0,640 (1,120)	-3,254 (2,577)	-1,442 (1,884)	-1,812 (1,609)
Norme sociale	Année 2014	**-3,087* (1,804)**	-1,039 (1,348)	**-2,048* (1,057)**	1,347 (1,605)	0,253 (1,099)	1,094 (1,121)	**-4,434* (2,578)**	-1,291 (1,885)	**-3,142* (1,610)**
Nombre d'observations		154952	154952	154952	154952	154952	154952	154952	154952	154952

Note : La moyenne de l'indicateur est inscrite pour le groupe témoin pour l'année 2014. Les écarts à ces valeurs sont inscrits pour les groupes traités. Les erreurs-types sont entre parenthèses. *** : p-value inférieure à 0,01 ; ** : p-value inférieure à 0,05 ; * : p-value inférieure à 0,10.

VII. Interprétation et discussion des résultats

L'expérimentation révèle qu'il y a eu, à la suite de l'envoi des messages, des réactions comportementales des entreprises en termes de déclaration sociale. Par conséquent, dans le cadre de la contribution sociale, il semble exister des arbitrages individuels qui interagissent avec le contenu des messages, les individus modifient leurs choix en fonction des informations dont ils disposent et des interprétations qu'ils en font.

1. Communiquer au sujet du dispositif de vérification de la fraude permet de modifier le comportement des contribuables

Ainsi, l'envoi d'informations aux contribuables à propos du dispositif de vérification de la déclaration sociale peut devenir un outil de politique publique permettant d'atteindre une amélioration du consentement à la contribution sociale. Le dispositif communicationnel destiné à expliquer l'architecture des règles en vigueur est donc aussi judicieux qu'un outil de coercition réservé à punir les comportements frauduleux. Cependant, il ne joue pas le même rôle puisque le premier agit comme un dispositif de prévention de la fraude, c'est-à-dire qu'il empêche le déclenchement de la fraude, tandis que le second se comporte comme un dispositif de sanction de la fraude, c'est-à-dire qu'il réprime le déclenchement de la fraude.

Dans cette campagne d'information expérimentale, nous souhaitions déterminer si la prise en compte des paramètres canoniques des modèles de fraude au prélèvement fiscal déterminait aussi le consentement à la contribution sociale. En effet, bien qu'il n'existe pas de déterminant théorique à la contribution sociale, son architecture institutionnelle ressemble à celle du système fiscal : il existe des dispositifs de vérification des déclarations sociales, de recouvrement des cotisations impayées et d'amendes en cas de constatation de fraude. Nous avons montré que les messages décrivant les sanctions financières modifiaient les comportements de déclaration d'embauche en CDI et que les messages expliquant les sanctions non financières modifiaient le comportement de recrutement en CDD. La prise en considération du risque de vérification des informations sociales a aussi conduit à mieux déclarer les embauches en CDD de courte durée. Ainsi dans une certaine mesure, ces effets démontrent que les paramètres incitatifs du consentement au prélèvement fiscal conditionnent le choix de contribuer à la protection sociale.
Nous souhaitions également établir si les paramètres enrichis d'interactions sociales des modèles étendus de fraude au prélèvement jouaient un rôle dans le consentement à la contribution sociale. Nous avons montré que le message expliquant la finalité des cotisations permettait d'accroître la déclaration d'embauche en CDI et que les messages insistant sur la notion de vol et de norme sociale réduisaient la création d'emploi en CDD. Ainsi il existe bien des considérations non

financières dans les choix de consentir à déclarer correctement ses informations sociales notamment en termes d'emploi.
Comparativement, le nombre de réactions comportementales liées à l'envoi des messages contenant des informations qui tiennent compte de la dimension morale de la contribution sociale est plus important que le nombre de réactions consécutives à l'envoi des messages contenant des informations sur l'architecture du contrôle de la fraude. Il semble donc que les interactions sociales soient des déterminants plus importants du consentement au prélèvement social que les déterminants financiers.

Les messages ont davantage modifié les arbitrages des entreprises en termes de déclaration d'emploi qu'en termes de déclaration de rémunération. Nous expliquons ce résultat par la nature des informations contenues dans les courriels. En effet, les messages reflétaient davantage la dimension de la réduction du travail dissimulé que la réduction de la fraude au prélèvement social. Ainsi, les organisations ont réagi en se conformant à leurs obligations déclaratives en termes d'emploi et non en termes de rémunérations ou de cotisations sociales. C'est un résultat étonnant, car les entreprises qui se conforment aux règles sur un versant de leurs obligations déclaratives sont susceptibles de frauder sur l'autre versant. Compte tenu de ce résultat, l'organisme de recouvrement des cotisations sociales sera contraint d'affiner ses futures campagnes d'information afin qu'elles contiennent des messages portant sur les deux versants de la fraude sociale.

Cependant, la plupart des différences relevées sur les indicateurs de déclaration sociale en stock après la diffusion de la campagne d'information ont été atténuées lorsque nous prenions en compte les observations antérieures au traitement. Ainsi, plusieurs groupes d'entreprises avaient des comportements différents de celui du groupe témoin, destinataire d'un message neutre, ce qui a invalidé des différences postérieures. De plus, nous ne pouvons pas rejeter l'hypothèse de nullité des écarts en double différence parce que, d'une part, la taille des effets est faible et que d'autre part, les estimations des écarts-types sont élevées. Par conséquent, l'effet pur des courriels n'a pas eu d'impact important en moyenne sur le comportement de déclaration sociale.

Néanmoins, il semble légitime de conserver les augmentations de déclaration d'embauche en CDI à la suite des messages décrivant les sanctions financières et la finalité des contributions sociales. En effet, ces deux groupes traités avaient un comportement similaire au groupe témoin avant l'expérimentation. Ces effets mêmes minimes témoignent d'un engagement plus important des organisations afin de mieux déclarer leurs futurs salariés en CDI. Les employeurs ont réagi pour

montrer à l'administration que les embauches se faisaient dans les règles. Les courriels ont donc amélioré le comportement des employeurs à s'engager dans une relation d'emploi durable.

2. L'effet principal de l'expérimentation : la réduction de la création d'emploi temporaire

Ce que nous retenons comme l'effet principal de l'expérimentation est la diminution de la création d'emploi en CDD. En terme contractuel, dans la relation de recouvrement, il semble qu'il y ait eu une diminution de l'aléa moral sur la déclaration d'employés à des conditions contraires à la réglementation. Cet effet s'observe pour trois groupes d'établissements : ceux qui ont été destinataires du message expliquant les sanctions non financières consécutives à la détection de la fraude ; ceux qui ont été destinataires du message insistant sur l'assimilation de la fraude au vol ; ceux qui ont été destinataires du message qui exprime l'idée que la norme sociale est de contribuer au financement de la protection sociale. Alors que le groupe témoin augmentait la création d'emploi en CDD, ces trois groupes traités ont réduit significativement leur embauche et ont été dissuadés de recourir au CDD.

Ce résultat est intéressant parce qu'il permet de reconsidérer les démarches de recrutement des entreprises. Le recrutement peut se définir comme la démarche de la prospection d'un futur employé, c'est-à-dire comme l'ensemble des stratégies mises en œuvre par l'employeur pour rechercher un candidat qui remplirait les conditions requises en termes de compétences et de qualifications pour un poste de travail vacant. Tandis que l'embauche est le résultat de la prospection. Par conséquent, il est judicieux de comprendre dans quelle mesure la communication ciblée au sujet du dispositif de vérification de la fraude sociale et ces trois variations de contenu de message ont pu contraindre les entreprises dans leur processus de recrutement en CDD.

D'un point de vue théorique, le recrutement est le versant employeur des modèles de recherche d'emploi (Barron et al., 1985). Le recrutement dépend du nombre de candidats potentiels et du coût de la recherche : il y a donc un ajustement par les prix ou par les quantités. Il existerait ainsi trois déterminants de la décision d'embaucher qui participent du processus de recrutement (Oyer et Schafer, 2014).

D'une part, le recrutement dépend de l'appariement entre les attentes et les préférences des employeurs en termes de productivité du travail et les offres de productivité des employés potentiels disponibles sur le marché du travail. Dans cette dimension, plus le différentiel entre les besoins et les opportunités est élevé, plus l'employeur devra prendre en considération des coûts de formation élevés (Baron et al., 1995). L'appariement s'effectue dès lors que le surplus dégagé dans la production est supérieur aux opportunités extérieures sur le marché. Plus l'appariement

est de qualité au début de l'embauche et plus la relation d'emploi est durable. En outre, plus l'appariement se réalise rapidement au début du processus de recrutement, plus la relation d'emploi est durable.

D'autre part, le recrutement dépend des coûts de recherche pour l'employeur, c'est-à-dire la difficulté de chercher et de trouver des employés pour l'organisation. Ces coûts dépendent du temps d'activation des différents canaux de prospection, de la quantité de candidats disponibles sur le marché et de la durée de vacance du poste à pourvoir (Rieucau et Marchal, 2010).

Enfin, la décision d'embaucher dépend de l'importance de l'asymétrie d'information, c'est-à-dire la capacité pour l'employeur de connaître les compétences des futurs employés avant de les recruter. Cette asymétrie dépend de la quantité et de la qualité des informations des intermédiaires sur le marché du travail : les curriculums vitae, les entretiens d'embauche et les opérateurs de placement, publics ou privés qui disposent de plateformes où circulent les offres et les demandes d'emploi.

Ces trois déterminants sont interdépendants les uns des autres : ainsi plus les coûts de recherche sont réduits plus l'asymétrie d'information est réduite et plus l'appariement sera rapide et de qualité. Il y a un arbitrage des employeurs entre la vitesse et la qualité du recrutement.

Les courriels qui étaient focalisés sur la fraude sociale auraient ainsi modifié les décisions des employeurs en renchérissant le coût du recrutement et en diminuant le nombre de candidats potentiels pour pourvoir les emplois en CDD.

Alors que les messages n'explicitaient pas directement le recours réglementé au contrat non permanent, les employeurs ont tenu compte de l'avertissement qui leur a été notifié s'il recourait au travail dissimulé[10]. Ainsi, on peut considérer que lorsqu'un employeur recourait au CDD pour un motif différent des motifs réglementaires, il a modifié son comportement de recrutement en optant pour moins de pratiques considérées comme illégales. On peut ainsi supposer que ces messages ont joué un rôle de réduction du comportement de dissimulation du travail : les

[10] C'est un contrat de travail atypique qui répond à plusieurs critères de recours. Ce contrat ne peut avoir ni pour effet ni pour objet de procurer durablement un emploi lié à l'activité « normale et permanente » de l'entreprise. Il ne peut être conclu que pour l'exécution d'une tâche précise et temporaire dans les cas suivants prévus par le Code du travail (avant la réforme de ce dernier le 22 septembre 2017) : le remplacement d'un salarié en cas d'absence, de suspension de son contrat de travail, de départ définitif précédant la suppression de son poste de travail ou d'attente de l'entrée en service effective d'un salarié recruté sous CDI ; l'accroissement temporaire de l'activité de l'entreprise ou travail saisonnier ; le « CDD d'usage » dans certains secteurs d'activité, du fait de l'activité exercée par l'entreprise et de l'emploi du salarié (la liste des secteurs concernés est fixée par décret ou par voie de convention ou d'accord collectif étendu) ; les contrats aidés (contrat unique d'insertion, contrat de professionnalisation, etc.) ; « CDD à objet défini » (créé en 2008) pour recruter des ingénieurs et des cadres en vue de la réalisation d'un objet défini si un accord de branche étendu ou, à défaut, un accord d'entreprise le prévoit. De plus, la durée du contrat et la date de fin de l'embauche sont explicitement mentionnées dans le contrat et son renouvellement ne peut excéder deux fois consécutives. La durée maximale, renouvellements compris, est portée jusqu'à 24 mois pour certains motifs (contrat exécuté à l'étranger, commande exceptionnelle à l'exportation, remplacement d'un salarié dont le départ définitif précède la suppression du poste) ou réduite à 9 mois pour d'autres (attente de l'entrée en service d'un salarié recruté en CDI, réalisation de travaux urgents par mesure de sécurité).

employeurs se sont conformés à la règle de recourir aux contrats dans le cadre réglementaire qui s'impose.

D'une part, il est possible que les recruteurs aient abandonné l'idée de recourir à des renouvellements de contrat pour les employés déjà présents dans l'établissement qui dépassaient les seuils légaux. Les employeurs ont ainsi craint la requalification de leur relation contractuelle avec des employés actuellement en CDD, à des conditions contraires à la réglementation, en CDI. Ces non-renouvellements de contrat se sont également accompagnés d'une mise en conformité avec les règles de fin de contrat. Étant donné qu'il n'y a pas eu plus de destruction d'emploi en CDD pour les trois groupes traités analysés, les fins de contrat ont eu lieu dans les mêmes conditions que le groupe témoin. Néanmoins, la prise en compte du non-renouvellement a conduit les employeurs à calculer des coûts de rupture de contrat plus importants, notamment en prenant en compte l'indemnité de fin de contrat. Celle-ci est égale au minimum à 10 % de la rémunération brute totale versée pendant la durée totale du contrat. Elle peut être limitée à 6 % par une convention ou un accord collectif de branche étendu, par convention ou accord d'entreprise ou d'établissement. Dans ce cas, des contreparties doivent être offertes au salarié, par exemple sous la forme d'un accès privilégié à la formation professionnelle (action de formation, bilan de compétences). Il existe ainsi toute une gamme de coûts de rupture qui ont été davantage pris en considération par ces employeurs. Par conséquent, cela a contraint à revoir à la baisse les possibilités de recrutement pour remplacer ces départs.

D'autre part, ils ont pu différer le recrutement de nouveaux salariés qui auraient pu être embauchés en CDD, mais qui ne rentraient pas dans les conditions d'embauche prévues par la loi. Les courriers ont ainsi réduit le *pool* d'individus susceptibles d'être embauchés à des conditions en adéquation avec la réglementation en vigueur. Ces employeurs ont donc prévu que les futurs embauchés devaient être formés à leur poste de travail, alors que s'ils avaient recouru à des renouvellements, les coûts de formation auraient été davantage internalisés.

En définitive, d'une façon ou d'une autre, les messages ont augmenté le coût d'opportunité de l'embauche des employés en CDD. Ces coûts entrent directement dans la demande de travail des employeurs. Les messages ont ainsi réduit l'adéquation entre une demande de travail pour un certain niveau de productivité formelle à un coût d'embauche faible et une offre de travail pour une même productivité formelle à des prétentions salariales plus élevées.

VIII. Conclusion

1. Une contribution à l'analyse de la relation de recouvrement des contributions sociales

Nous souhaitions caractériser dans quelle mesure une campagne d'information, mobilisant plusieurs types de messages qui contiennent des renseignements sur l'architecture du système de vérification de la déclaration sociale, les modalités de sanctions, et l'utilité de la contribution sociale, permettait d'augmenter le consentement au prélèvement social et de réduire la fraude sociale des entreprises.

Pour y répondre, nous avons conduit une expérimentation, dont l'objectif était de tester l'impact de onze contenus de messages différents, qui couvraient à la fois les paramètres traditionnels de la technologie de contrôle de la fraude et des paramètres enrichis d'interactions sociales sur les indicateurs de déclarations sociales en stock et sur les flux d'emploi entre 2013 et 2014. Nous avons montré que les principales réactions comportementales concernaient les pratiques de recrutement en contrat à durée déterminée. Les messages insistant sur l'idée que la fraude c'est du vol, que la norme sociale est de consentir au prélèvement et que la fraude occasionnait des sanctions non financières ont davantage réduit les flux de création d'emploi non permanent.

Au regard de nos éléments de discussion, l'expérimentation a montré les bénéfices d'un dispositif préventif dans l'éventail des possibilités pour l'agence de recouvrement des cotisations sociales afin de réduire la fraude. Elle a permis de mettre en lumière les réactions comportementales des entreprises face aux informations contenues dans les courriels. D'un point de vue organisationnel et dans une perspective de réduction de la fraude sociale, ces réactions comportementales donnent à voir une réduction de l'aléa moral sur les obligations déclaratives qui se traduit par un meilleur consentement pour les employeurs à se conformer aux règles du recouvrement.
Toutefois, cette campagne d'information a davantage permis d'augmenter la dimension déclarative sur l'effectif que de réduire la fraude aux cotisations sociales. La prévention a donc joué un rôle positif sur la prise en compte de la réglementation sur le travail dissimulé par les employeurs, c'est-à-dire qu'ils ont modifié leurs modes de recrutement afin de ne pas être pénalisés si l'administration venait à vérifier la composition de la main-d'œuvre.
De plus, les messages contenant des informations sur les interactions sociales liées à la contribution sociale ont engendré davantage d'effets que les messages contenant des informations liées aux paramètres canoniques des modèles de fraude aux prélèvements obligatoires. Ce résultat permet de renforcer la spécificité de la contribution sociale par rapport

au prélèvement fiscal. Il renforce également l'idée selon laquelle les entreprises possèdent une moralité contributive au même titre que la moralité fiscale (Alm et McClellan, 2012).
L'explication des conditions de vérification de la déclaration sociale associée à des messages destinés à réaliser des arbitrages contenant davantage de considérations morales envers la contribution sociale permet de modifier les comportements de déclaration sociale. Dans ce contexte, nous fournissons un déterminant supplémentaire à un paramétrage efficient du dispositif de vérification de l'information. Lorsque le principal explique le fonctionnement du contrôle et qu'il associe cette explication à des considérations qui sont pertinentes dans les décisions d'effectuer ou non une action opportuniste, alors l'explication permet de réduire l'aléa moral. Elle permet de légitimer l'existence du dispositif de vérification de l'information. Les intérêts des deux parties de la relation d'agence peuvent ainsi converger.

Malgré ces résultats, nous n'observons finalement qu'un faible nombre de réactions par rapport au nombre et à la nature des indicateurs employés pour mesurer les effets. Ce constat peut paraître dérisoire compte tenu du nombre de messages envoyés. Cependant, il se trouve que les précédentes expérimentations de terrain conduites sur les entreprises ont également montré de faibles réactions comportementales à la suite de l'envoi de messages informatifs. Par conséquent, cela indique que l'information ciblée ne doit pas être le seul outil dont les autorités disposent pour prévenir la fraude.
Il faut donc concevoir cette expérimentation comme une première étape exploratoire vers un approfondissement des dispositifs de lutte contre la fraude sociale. L'outil de prévention doit ainsi être mobilisé après une réflexion sur le choix des informations à communiquer et sur le résultat escompté. Si la campagne était destinée à réduire la fraude sociale, elle n'a finalement permis que de réduire l'une des composantes de la fraude, à savoir la déclaration de l'effectif. Ce changement comportemental permet de penser qu'il existe des marges de manœuvre pour l'Acoss afin de réduire le travail dissimulé, sans forcément qu'il y ait un gain en termes de consentement au prélèvement social. À cause de cet effet ambivalent, le résultat final paraît peu avantageux pour l'Acoss qui cherche à augmenter les recettes de cotisations sociales, tout du moins à éviter qu'une partie des contributions ne soit pas recouvrée. Une autre campagne d'information mieux calibrée sur l'aspect contributif devrait ainsi permettre de réduire les deux versants de la fraude sociale.

2. Les limites à la portée des résultats

Le faible nombre de résultats pose la question de la crédibilité des informations transmises et de la manière dont elles ont été perçues par les destinataires. Ainsi, notre analyse manque d'un regard plus qualitatif sur les effets observés. En l'absence de ce type d'approche, il est difficile de déterminer parmi les motivations du recours au CDD (un besoin limité dans le temps, une

évaluation des compétences avant une embauche durable [période d'essai étendue], une réduction du risque si l'activité est ralentie, à une habitude de recrutement, un contournement de la réglementation des CDI…), laquelle serait la plus susceptible d'être modifiée à la suite de l'envoi des courriers. Si l'effet majeur des messages tend à être confirmé, alors cela encouragerait à créer une recherche plus poussée sur les déterminants du recours au CDD.

Un travail approfondi de terrain sur l'appréciation que font les employeurs sur leurs obligations déclaratives et sur les manières d'employer la main-d'œuvre en emploi temporaire permettrait de mettre en contraste les interprétations que nous faisons sur les indicateurs de flux d'emploi. Un protocole d'enquête dont l'objet porterait sur les modes de gestion de la déclaration sociale ferait ressortir des aspects qui n'étaient pas pris en compte initialement dans le protocole expérimental.

De plus, nos interprétations doivent être considérées au conditionnel en l'absence de mesures sur la nature des contrats dans les établissements, le contenu des tâches réalisées, et les multiples agencements dans les relations d'emploi (Kramarz et Michaud, 2010). Il est certain que d'autres déterminants du recrutement ont été pris en considération dans l'ajustement de la masse salariale dans les groupes d'établissements qui ont réduit la création d'emploi en CDD.

Plus globalement, un programme de recherche qui porterait sur la teneur des réactions des employeurs à la suite de la réception des messages permettrait d'affiner les interprétations : comment est interprété le message ? Quel est le premier acteur qui le réceptionne ? Quelles actions sont entreprises ?

La motivation pour observer ces comportements doit être conditionnée à des précautions importantes, puisque la recherche aurait comme ambition de détecter *in situ* la fraude sociale.

De plus, une question peut se poser quant à la durabilité des effets observés. La plupart des différences entre le groupe témoin et les groupes traités ne s'observent que lors d'un seul trimestre. Dès lors, il est difficile de juger si l'explication du dispositif de contrôle a permis de durablement changer les comportements des établissements fraudeurs.

Comme nous l'avons précisé à plusieurs reprises au cours de ce chapitre, le manque de support théorique propre à la contribution sociale concourt également à une appréciation moins précise des effets mesurés. L'assurance sociale n'est pas un dispositif fiscal comme les autres parce qu'elle ouvre un droit à prestations et à protection. De plus, l'analyse des déterminants de la demande de travail illégal doit être approfondie en tenant compte des interactions contractuelles avec l'agence de recouvrement des cotisations. Par conséquent, nos résultats souffrent d'un manque de prise de recul théorique nécessaire pour mieux appréhender les observations.

3. Des pistes ouvertes pour approfondir le rôle du dispositif de vérification de la déclaration sociale dans le consentement à la contribution sociale

Les futurs travaux appliqués et théoriques pourront prendre en considération ces limites afin de concevoir des designs expérimentaux plus aboutis et des modélisations plus avancées et propres à la contribution sociale tout en tenant compte de la dimension relationnelle et du rôle joué par le dispositif de vérification de l'information.

Nous avons considéré des indicateurs moyens, sans étudier la distribution des comportements frauduleux parmi les établissements traités. Or, la prévalence de la fraude, c'est-à-dire le nombre réel d'entreprises réalisant de mauvaises déclarations sociales, peut être faible, tandis que la proportion d'entreprises intègres est plus importante. Par conséquent, il est possible que les effets que nous mesurons ne permettent pas de visualiser les changements parmi les seules entreprises fraudeuses.

Ainsi, nous avons pu produire une sous-estimation de l'ampleur des modifications comportementales à l'œuvre. En effet, dans la stratégie d'identification, nous n'avons pas choisi de distinguer les entreprises fraudeuses des entreprises intègres. Or, dans l'ensemble des établissements, on peut distinguer celles qui ont réagi positivement au traitement, en se conformant mieux aux règles, et celles qui ont continué d'agir illégalement. Cela suppose d'être en mesure d'appliquer les techniques de détection des entreprises fraudeuses pour être à même de conduire une analyse différenciée entre ces groupes (Joubert, 2009). Une façon de procéder serait d'utiliser les informations issues des vérifications opérées par les Urssaf pour qualifier les différents types d'entreprises entre celles qui fraudent et celles qui ne fraudent pas.

Un autre moyen d'approfondir ce sujet serait de définir des indicateurs de déclarations sociales complémentaires à ceux qui sont employés dans cette expérimentation. Une première voie de mesure serait de considérer les différents types de rémunérations déclarées (salaire, primes, participation au bénéfice, épargne salariale, compte épargne temps, etc.) ainsi que les déclarations de cotisation, plutôt que les déclarations d'assiette.

Le traitement de l'explication du dispositif de vérification pourrait également être modifié pour inclure les inspecteurs des Urssaf. Dans cette optique, leurs activités de contrôle seraient susceptibles d'être complétées par des activités de prévention de la fraude, au moyen de formation ou d'accompagnement des employeurs. Plutôt que d'intervenir dans les entreprises pour vérifier les comptes, il s'agirait d'expérimenter des interventions physiques pour inculquer les règles, les risques et les normes sociales de la contribution.

Une voie d'étude complémentaire serait de faire varier la temporalité de l'explication du dispositif, c'est-à-dire de choisir des périodes de la relation contractuelle pour révéler l'existence du contrôle. Il est en effet possible d'établir que l'une des conditions de la faible détection de réactions à la suite des messages est liée au fait que les employeurs connaissaient déjà l'architecture du contrôle. Restreindre l'étude aux entreprises nouvellement créées permettrait d'approfondir l'effet de l'explicitation du dispositif de vérification de la déclaration sociale.

Une autre piste de réflexion concerne le choc de simplification des obligations déclaratives survenu depuis 2016 : il s'agit du passage à la DSN. Elle constitue la transmission mensuelle d'un fichier unique contenant l'ensemble des déclarations sociales à un unique interlocuteur. Cet interlocuteur communique aux organismes de protection sociale, qui collectaient précédemment les notices déclaratives, les informations dont ils souhaitent disposer. Elle a remplacé progressivement une trentaine de déclarations périodiques et événementielles dont l'attestation de salaire pour le paiement des indemnités journalières en cas d'arrêt maladie, de congés maternité, paternité, l'attestation destinée à Pôle emploi en cas de fin de contrat de travail, les déclarations et enquêtes de mouvements de main-d'œuvre et la déclaration unifiée des cotisations sociales, le bordereau de cotisations Urssaf. En 2017, la DADS a été complètement intégrée à ce dispositif. De plus, elle s'effectue de manière dématérialisée. Par conséquent, elle permet de simplifier les démarches des déclarants, réduit le risque d'erreurs, et réduit le nombre de vérifications administratives.
Comme le notaient Bailey et Turner (2001), la simplification des démarches administratives constitue un levier important de la réduction de la fraude à la contribution sociale. Il serait donc judicieux de déterminer dans quelle mesure la DSN a permis de réduire la quantité de mauvaises déclarations sociales et d'améliorer la qualité de la déclaration et si elle engendre un meilleur consentement à la contribution sociale. Dans ce contexte, il s'agit de déterminer si le fait de réduire le nombre de déclarations sociales à vérifier permet d'assurer une meilleure coordination de l'activité de recouvrement. La recherche porterait ainsi spécifiquement sur les effets du dispositif de vérification à l'issue du choc de simplification déclarative.

Enfin, une dernière voie de recherche serait de pouvoir répliquer ce traitement consistant à expliquer le rôle du dispositif de vérification à d'autres relations d'agence. Est-ce un moyen qui serait utile seulement dans des relations où les comportements frauduleux sont susceptibles d'exister, comme le piratage numérique des œuvres audiovisuelles ? De quelle manière l'explication de la vérification de l'information peut-elle s'effectuer pour qu'elle soit optimale ? Comment tenir compte des effets ambivalents du contrôle dans un cadre d'explication ?

Annexes du chapitre 1

4. Contenu des messages envoyés aux groupes traités et au groupe témoin

Les différents contenus des messages adressés aux entreprises sont renseignés dans le tableau A1. Toutes les entreprises ont reçu un courriel identique à celui du groupe témoin. Le traitement consistait à rédiger un paragraphe spécifique après le rappel des règles de déclaration sociale. Le paraphe modifié est rédigé en gras dans le tableau ci-dessous. Dans ce contexte, les entreprises recevaient une information générale sur les règles du contrôle de la déclaration sociale et un message spécifique sur les paramètres du contrôle ou sur des considérations morales.

Tableau A1 : Contenu des courriels adressés aux entreprises cotisantes

Groupe Message testé	Contenu
Témoin (1) Neutre	Objet —L'Urssaf vous informe : le point sur les obligations des employeurs. Madame, Monsieur, La lutte contre les fraudes aux cotisations et contributions sociales, et en particulier contre le travail dissimulé, constitue un enjeu majeur. Elle vise à garantir la reconnaissance des droits des salariés et une saine concurrence entre les entreprises. Il y a dissimulation d'emploi salarié en cas : • d'absence de Déclaration préalable à l'embauche (DPAE) à l'Urssaf préalablement à tout emploi ; • de non remise d'un bulletin de salaire ou d'établissement d'un bulletin de salaire ne comportant pas le nombre d'heures réellement effectuées ; • d'utilisation d'un faux statut (exemple : recours infondé au statut de travailleur indépendant) ; • de non accomplissement des déclarations relatives aux salaires ou aux cotisations sociales. Les Urssaf et l'ensemble de leurs partenaires sont engagés très largement et quotidiennement dans de multiples actions visant à garantir le respect de la législation sociale.

Groupe Message testé	Contenu
Traité (2) Risque de détection	Objet —L'Urssaf vous informe : le point sur les obligations des employeurs Madame, Monsieur, La lutte contre les fraudes aux cotisations et contributions sociales, et en particulier contre le travail dissimulé, constitue un enjeu majeur. Elle vise à garantir la reconnaissance des droits des salariés et une saine concurrence entre les entreprises. Il y a dissimulation d'emploi salarié en cas : • d'absence de Déclaration préalable à l'embauche (DPAE) à l'Urssaf préalablement à tout emploi ; • de non remise d'un bulletin de salaire ou d'établissement d'un bulletin de salaire ne comportant pas le nombre d'heures réellement effectuées ; • d'utilisation d'un faux statut (exemple : recours infondé au statut de travailleur indépendant) ; • de non accomplissement des déclarations relatives aux salaires ou aux cotisations sociales. **Les progrès réalisés dans la détection de la fraude sociale sont importants, avec notamment le développement des actions de contrôle ciblées, la coopération renforcée entre l'ensemble des acteurs de la lutte contre la fraude et la mobilisation d'outils statistiques sophistiqués. Les services de contrôle vérifient de plus en plus aisément si les obligations déclaratives ont bien été respectées. Les redressements de cotisations opérés par les Urssaf dans le cadre de la lutte contre le travail dissimulé ont ainsi doublé ces quatre dernières années.**
Traité (3) Sanctions monétaires	Objet —L'Urssaf vous informe : le point sur les obligations des employeurs Madame, Monsieur, La lutte contre les fraudes aux cotisations et contributions sociales, et en particulier contre le travail dissimulé, constitue un enjeu majeur. Elle vise à garantir la reconnaissance des droits des salariés et une saine concurrence entre les entreprises. Il y a dissimulation d'emploi salarié en cas : • d'absence de Déclaration préalable à l'embauche (DPAE) à l'Urssaf préalablement à tout emploi ; • de non remise d'un bulletin de salaire ou d'établissement d'un bulletin de salaire ne comportant pas le nombre d'heures réellement effectuées ; • d'utilisation d'un faux statut (exemple : recours infondé au statut de travailleur indépendant) ; • de non accomplissement des déclarations relatives aux salaires ou aux cotisations sociales. **Les employeurs en infraction prennent des risques financiers importants. En cas de travail dissimulé, l'employeur doit verser les sommes dues, ainsi que, le cas échéant, le remboursement rétroactif des exonérations et des allègements de cotisations sociales dont il aurait bénéficié. Enfin, il s'expose à des amendes qui peuvent atteindre 45 000 € pour une personne physique et 225 000 € pour une personne morale.**

Groupe Message testé	Contenu
Traité (4) Sanctions non monétaires	Objet —L'Urssaf vous informe : le point sur les obligations des employeurs Madame, Monsieur, La lutte contre les fraudes aux cotisations et contributions sociales, et en particulier contre le travail dissimulé, constitue un enjeu majeur. Elle vise à garantir la reconnaissance des droits des salariés et une saine concurrence entre les entreprises. Il y a dissimulation d'emploi salarié en cas : • d'absence de Déclaration préalable à l'embauche (DPAE) à l'Urssaf préalablement à tout emploi ; • de non remise d'un bulletin de salaire ou d'établissement d'un bulletin de salaire ne comportant pas le nombre d'heures réellement effectuées ; • d'utilisation d'un faux statut (exemple : recours infondé au statut de travailleur indépendant) ; • de non accomplissement des déclarations relatives aux salaires ou aux cotisations sociales. **Les employeurs en infraction s'exposent à des sanctions pénales importantes. Ainsi, le tribunal correctionnel peut prononcer une peine d'emprisonnement allant jusqu'à trois ans. En outre, des peines complémentaires peuvent être prononcées telles que l'interdiction d'exercer son activité, la dissolution et la fermeture de l'établissement et la confiscation du matériel et des stocks.**
Traité (5) Coûts cachés	Objet —L'Urssaf vous informe : le point sur les obligations des employeurs Madame, Monsieur, La lutte contre les fraudes aux cotisations et contributions sociales, et en particulier contre le travail dissimulé, constitue un enjeu majeur. Elle vise à garantir la reconnaissance des droits des salariés et une saine concurrence entre les entreprises. Il y a dissimulation d'emploi salarié en cas : • d'absence de Déclaration préalable à l'embauche (DPAE) à l'Urssaf préalablement à tout emploi ; • de non remise d'un bulletin de salaire ou d'établissement d'un bulletin de salaire ne comportant pas le nombre d'heures réellement effectuées ; • d'utilisation d'un faux statut (exemple : recours infondé au statut de travailleur indépendant) ; • de non accomplissement des déclarations relatives aux salaires ou aux cotisations sociales. **Pour une entreprise, les sanctions financières ne se limitent pas aux cas de travail dissimulé de ses propres employés. Elles s'appliquent aussi en cas de recours à une entreprise sous-traitante qui ne déclare pas son personnel. Serait ainsi à acquitter par le donneur d'ordre, au titre de la solidarité financière, le montant des cotisations non versées par son sous-traitant. De plus aucune garantie en cas de malfaçon des travaux réalisés ne serait applicable.**

Groupe Message testé	Contenu
Traité (6) Coûts différés	Objet —L'Urssaf vous informe : le point sur les obligations des employeurs Madame, Monsieur, La lutte contre les fraudes aux cotisations et contributions sociales, et en particulier contre le travail dissimulé, constitue un enjeu majeur. Elle vise à garantir la reconnaissance des droits des salariés et une saine concurrence entre les entreprises. Il y a dissimulation d'emploi salarié en cas : • d'absence de Déclaration préalable à l'embauche (DPAE) à l'Urssaf préalablement à tout emploi ; • de non remise d'un bulletin de salaire ou d'établissement d'un bulletin de salaire ne comportant pas le nombre d'heures réellement effectuées ; • d'utilisation d'un faux statut (exemple : recours infondé au statut de travailleur indépendant) ; • de non accomplissement des déclarations relatives aux salaires ou aux cotisations sociales. **Le constat de travail dissimulé peut fragiliser durablement l'activité des entreprises contrevenantes. Les sanctions en cas de travail dissimulé s'appliquent aux 5 années civiles qui précèdent la constatation de l'infraction et à l'année en cours. De plus, les sanctions sont doublées en cas de récidive. En outre, l'entreprise s'expose à une interdiction de marché public pendant 5 ans. Enfin, seront également refusées les aides à l'emploi et à la formation professionnelle pendant 5 ans.**
Traité (7) Contribution production bien public	Objet —L'Urssaf vous informe : le point sur les obligations des employeurs Madame, Monsieur, La lutte contre les fraudes aux cotisations et contributions sociales, et en particulier contre le travail dissimulé, constitue un enjeu majeur. Elle vise à garantir la reconnaissance des droits des salariés et une saine concurrence entre les entreprises. Il y a dissimulation d'emploi salarié en cas : • d'absence de Déclaration préalable à l'embauche (DPAE) à l'Urssaf préalablement à tout emploi ; • de non remise d'un bulletin de salaire ou d'établissement d'un bulletin de salaire ne comportant pas le nombre d'heures réellement effectuées ; • d'utilisation d'un faux statut (exemple : recours infondé au statut de travailleur indépendant) ; • de non accomplissement des déclarations relatives aux salaires ou aux cotisations sociales. **Le financement de la Sécurité sociale repose sur les cotisations et contributions des salariés et de l'ensemble des employeurs. L'acquittement des montants dus est le garant de sa pérennité.** **Sur 100 € de cotisations et contributions sociales :** • **37 € sont versés à la branche Maladie pour le financement des prestations maladie, maternité, et invalidité/décès,** • **23 € assurent le financement de la branche Vieillesse (pensions du régime de base, pensions de réversion et assurance veuvage),** • **17 € financent la branche Famille (allocations familiales, etc.),** • **3 € financent les prestations de la branche Accidents du travail et maladies professionnelles,** • **20 € sont versés à d'autres organismes (Fond de solidarité vieillesse, Caisse d'amortissement de la dette sociale, etc.), dont 7 € à l'Unedic au titre de l'assurance chômage.**

Groupe Message testé	Contenu
Traité (8) « Frauder c'est voler »	Objet —L'Urssaf vous informe : le point sur les obligations des employeurs Madame, Monsieur, La lutte contre les fraudes aux cotisations et contributions sociales, et en particulier contre le travail dissimulé, constitue un enjeu majeur. Elle vise à garantir la reconnaissance des droits des salariés et une saine concurrence entre les entreprises. Il y a dissimulation d'emploi salarié en cas : • d'absence de Déclaration préalable à l'embauche (DPAE) à l'Urssaf préalablement à tout emploi ; • de non remise d'un bulletin de salaire ou d'établissement d'un bulletin de salaire ne comportant pas le nombre d'heures réellement effectuées ; • d'utilisation d'un faux statut (exemple : recours infondé au statut de travailleur indépendant) ; • de non accomplissement des déclarations relatives aux salaires ou aux cotisations sociales. **La fraude constitue une atteinte au pacte social qui unit l'ensemble des citoyens. Elle porte atteinte à la solidarité nationale et fait obstacle à la cohésion sociale. Frauder c'est voler, et nous avons tous à y perdre. Celui qui fraude sera sanctionné.**
Traité (9) Publicité des noms des fraudeurs	Objet —L'Urssaf vous informe : le point sur les obligations des employeurs Madame, Monsieur, La lutte contre les fraudes aux cotisations et contributions sociales, et en particulier contre le travail dissimulé, constitue un enjeu majeur. Elle vise à garantir la reconnaissance des droits des salariés et une saine concurrence entre les entreprises. Il y a dissimulation d'emploi salarié en cas : • d'absence de Déclaration préalable à l'embauche (DPAE) à l'Urssaf préalablement à tout emploi ; • de non remise d'un bulletin de salaire ou d'établissement d'un bulletin de salaire ne comportant pas le nombre d'heures réellement effectuées ; • d'utilisation d'un faux statut (exemple : recours infondé au statut de travailleur indépendant) ; • de non accomplissement des déclarations relatives aux salaires ou aux cotisations sociales. **La condamnation pour travail dissimulé peut donner lieu à une publicité légale. En effet, un jugement prononcé par le tribunal correctionnel est susceptible à la demande de l'Urssaf d'une publication auprès de journaux spécialisés ou au fronton de l'établissement concerné. En cas de fraude sanctionnée par un tribunal, le personnel, les clients et les fournisseurs peuvent le cas échéant avoir accès à cette information.**

Groupe Message testé	Contenu
Traité (10) Mise en danger d'autrui	Objet —L'Urssaf vous informe : le point sur les obligations des employeurs Madame, Monsieur, La lutte contre les fraudes aux cotisations et contributions sociales, et en particulier contre le travail dissimulé, constitue un enjeu majeur. Elle vise à garantir la reconnaissance des droits des salariés et une saine concurrence entre les entreprises. Il y a dissimulation d'emploi salarié en cas : • d'absence de Déclaration préalable à l'embauche (DPAE) à l'Urssaf préalablement à tout emploi ; • de non remise d'un bulletin de salaire ou d'établissement d'un bulletin de salaire ne comportant pas le nombre d'heures réellement effectuées ; • d'utilisation d'un faux statut (exemple : recours infondé au statut de travailleur indépendant) ; • de non accomplissement des déclarations relatives aux salaires ou aux cotisations sociales. **La fraude met en danger des personnes. Les salariés dont l'activité est dissimulée ne peuvent bénéficier de l'ensemble des droits sociaux. Ils ne sont pas couverts au titre d'un accident du travail ou d'une maladie professionnelle. En cas de maladie, il ne pourrait pas bénéficier d'un revenu de remplacement et du remboursement de ses frais de soins. De plus, une personne dont l'activité est non déclarée ne voit pas ses droits à la retraite reconnus.**
Traité (11) Norme sociale	Objet —L'Urssaf vous informe : le point sur les obligations des employeurs Madame, Monsieur, La lutte contre les fraudes aux cotisations et contributions sociales, et en particulier contre le travail dissimulé, constitue un enjeu majeur. Elle vise à garantir la reconnaissance des droits des salariés et une saine concurrence entre les entreprises. Il y a dissimulation d'emploi salarié en cas : • d'absence de Déclaration préalable à l'embauche (DPAE) à l'Urssaf préalablement à tout emploi ; • de non remise d'un bulletin de salaire ou d'établissement d'un bulletin de salaire ne comportant pas le nombre d'heures réellement effectuées ; • d'utilisation d'un faux statut (exemple : recours infondé au statut de travailleur indépendant) ; • de non accomplissement des déclarations relatives aux salaires ou aux cotisations sociales. **La plupart des employeurs déclarent leurs salariés et acquittent les cotisations réellement dues. La fraude sociale concerne une minorité d'employeurs qui ne respectent pas les règles et sont sévèrement sanctionnés.**

5. Test d'équivalence entre l'échantillon initial et l'échantillon final

Le phénomène du multitraitement a contraint a opéré un traitement de l'échantillon pour identifier les établissements directement abonnés et cotisants à l'Acoss. Cet échantillon possède ainsi des caractéristiques différentes par rapport à l'échantillon initial. Il y a notamment une surreprésentation des établissements localisés en Île-de-France, une surreprésentation du secteur public, une sous-représentation des entrepreneurs individuels et une surreprésentation des personnes morales et organisme soumis au droit.

Tableau A2 : Test d'équivalence entre l'échantillon initial et l'échantillon final

	Échantillon initial	Différence avec l'échantillon final	Test de Wald
Nombre d'établissements	440000	77476	
Alsace	2,754	0,074 (0,064)	1,16
Aquitaine	5,564	**-0,492***** **(0,089)**	5,54
Auvergne	1,961	**-0,329***** **(0,053)**	6,16
Basse-Normandie	2,164	-0,094 (0,057)	1,66
Bourgogne	2,306	**0,247***** **(0,059)**	4,19
Bretagne	5,124	**-0,864***** **(0,085)**	10,18
Centre-Val de Loire	3,285	**0,614***** **(0,070)**	8,72
Champagne-Ardenne	1,717	**0,467***** **(0,052)**	9,05
Corse	0,938	**-0,359***** **(0,036)**	9,86
DOM	2,198	**1,068***** **(0,059)**	18,07
Franche-Comté	1,547	-0,065 (0,048)	1,36
Haute Normandie	2,256	**0,420***** **(0,059)**	7,16
Île-de-France	20,150	**5,574***** **(0,159)**	35,13
Languedoc-Roussillon	4,548	**-1,281***** **(0,080)**	16,11
Limousin	1,061	-0,018 (0,040)	0,45
Lorraine	2,896	**0,262***** **(0,066)**	3,98
Midi-Pyrénées	4,651	**-0,677***** **(0,081)**	8,34
Nord Pas de Calais	4,336	**-0,634***** **(0,079)**	8,08
Pays de la Loire	4,875	**0,266***** **(0,084)**	3,15
Picardie	2,183	**0,421***** **(0,058)**	7,29
Poitou-Charentes	2,568	0,094 (0,062)	1,53
PACA	9,606	**-2,659***** **(0,113)**	23,61
Rhône-Alpes	11,000	**-1,818***** **(0,121)**	15,07
Étranger	0,312	**-0,219***** **(0,021)**	10,66
Entrepreneur individuel	18,821	**-10,528***** **(0,147)**	71,5
Groupement de droit privé non doté de la personnalité morale	0,335	0,007 (0,023)	0,29
Personne morale de droit étranger	0,258	0,000	0,02

	Échantillon initial	Différence avec l'échantillon final	Test de Wald
		(0,020)	
Personne morale de droit public soumise au droit commercial	0,055	**0,144*** (0,011)**	13,4
Société commerciale	68,976	**-2,662*** (0,181)**	14,71
Autre personne morale immatriculée au registre du commerce et des sociétés (RCS)	1,549	**0,298*** (0,049)**	6,11
Personne morale et organisme soumis au droit administratif	2,484	**8,150*** (0,074)**	110,57
Organisme privé spécialisé	0,317	**0,232*** (0,023)**	10,06
Groupement de droit privé	7,205	**4,359*** (0,105)**	41,54
Agriculture sylviculture et pêche	0,090	**0,026*** (0,012)**	2,17
Industries extractives	0,073	**0,050*** (0,011)**	4,5
Industrie manufacturière	5,001	**0,603*** (0,086)**	7,04
Production et distribution d'électricité, de gaz, de vapeur et d'air conditionné	0,040	**0,021*** (0,008)**	2,62
Production et distribution d'eau ; assainissement, gestion des déchets et dépollution	0,315	**0,339*** (0,024)**	14,4
Construction	28,854	**-3,435*** (0,176)**	19,56
Commerce ; réparation d'automobiles et de motocycles	14,096	**-2,029*** (0,134)**	15,1
Transports et entreposage	1,764	**0,469*** (0,052)**	8,97
Hébergement et restauration	21,146	**-7,774*** (0,156)**	49,9
Information et communication	1,573	**0,484*** (0,050)**	9,77
Activités financières et d'assurance	2,140	**0,291*** (0,057)**	5,12
Activités immobilières	1,714	**-0,269*** (0,050)**	5,38
Activités spécialisées, scientifiques et techniques	5,285	**0,458*** (0,088)**	5,22
Activités de services administratifs et de soutien	5,080	**-0,942*** (0,084)**	11,16
Administration publique	1,619	**6,563*** (0,062)**	105,81
Enseignement	1,488	**0,861*** (0,049)**	17,52
Santé humaine et action sociale	3,632	**1,728*** (0,075)**	22,94
Arts, spectacles et activités récréatives	1,832	**2,065*** (0,056)**	36,61
Autres activités de services	4,246	**0,474*** (0,079)**	5,98
Activités extraterritoriales	0,011	**0,017*** (0,005)**	3,78

Note : les proportions de l'échantillon initial servent de valeur de référence. Les erreurs-types sont entre parenthèses. *** : p-value inférieure à 0,01 ; ** : p-value inférieure à 0,05 ; * : p-value inférieure à 0,10.

Chapitre 2.
Réduire l'intensité de la surveillance électronique des performances des salariés : une expérimentation dans un centre d'appels sous-traitant[11]

[11] Ce chapitre est issu d'un projet de recherche financé par le Club informatique des grandes entreprises françaises (Cigref) dans le cadre de l'appel à étude sur les changements organisationnels et l'innovation numérique à l'horizon 2020, encadré par Nathalie Greenan et auquel a participé Isabelle Gilet. Intitulé Centre d'Appels 2020, ce projet a été hébergé au Centre d'Études de l'Emploi et du Travail (ex-CEE). Il a fait l'objet d'une valorisation sous la forme d'un ouvrage de recherche publié aux éditions Lavoisier dans la collection « Espaces Numériques » (Gilet, Greenan et Le Gall, 2016). Les résultats produits par ce projet ont été présentés en 2015 aux 32ème Journées de Microéconomie Appliquée organisées par le CEE-M (ex-LAMETA) à Montpellier, à la 3ème Journée annuelle des doctorants de l'ERUDITE (Jade), au 50ème congrès de la Société d'Ergonomie de Langue Française (SELF) et à l'European Conference on Psychology and Behavioral Science 2015 (ECP2015). Enfin, le Connaissance de l'Emploi n°126 reprend les principaux résultats de cette étude.

I. Introduction

En 2008, une équipe de chercheurs en sciences de l'information du Massachusetts Institute of Technology (MIT) publiait une étude relative à l'utilisation de « badges sociométriques » dans une entreprise d'informatique pour montrer les liens existants entre les interactions sociales au travail et le niveau de productivité des salariés (Wu et al., 2008). Ces objets, traditionnellement utilisés pour pointer le temps de travail des employés, mesuraient le temps de conversation du personnel, la variation sonore des échanges, la proximité physique des interlocuteurs, les déplacements sur le lieu de travail ainsi que le type de communication, qu'elle soit physique ou électronique. Leur recherche a notamment montré que des communications directes en face à face augmentaient significativement la performance au travail y compris sur des tâches complexes.

Dix ans plus tard, ces objets sont commercialisés par la société Humanyze fondée par Ben Waber, l'un des chercheurs ayant participé à ce projet de recherche. Ces badges sont utilisés par plus d'une cinquantaine de moyennes et grandes entreprises aux États-Unis, mais également par la National Aeronautics and Space Administration (NASA). Ils sont l'avatar des « objets connectés » au travail : en collectant des données sur les comportements au travail des employés, ils permettent à ces entreprises de surveiller en temps réel l'activité et les interactions sociales. Avec l'avènement de la gestion de données en masse, le Big Data, il est ainsi désormais possible de pister les employés afin d'améliorer leur productivité. Les informations récoltées par les badges servent notamment à déterminer le niveau de collaboration d'un salarié avec ses pairs. Le rapport intitulé The Enterprise Wearables Journey (Gownder et al., 2014) notait ainsi l'ambivalence de ces outils de surveillance au travail : « le salarié se connecte au système pour comprendre comment améliorer sa manière de collaborer et pour se comparer à ses collègues, tandis que les managers ont de leur côté accès à des données agrégées pour faire évoluer la politique de l'entreprise sur le sujet si nécessaire ».

1. Les éléments de définition de la surveillance des performances au travail

La surveillance au travail peut se définir comme tout dispositif visant à collecter, enregistrer, traiter, rendre compte et partager de l'information sur le comportement des personnes travaillant au sein d'une organisation productive (Nebeker et Tatum, 1993 ; Alge et Hansen, 2014).

L'intensité de la surveillance est le degré d'observation des comportements. Trois axes permettent de circonscrire l'intensité : la quantité d'information qui doit être surveillée ; la temporalité de l'information qui doit être surveillée ; l'échelon individuel où l'information qui doit être surveillée. Une fois ces trois axes paramétrés, l'organisation peut alors déterminer le dispositif qui sera

adopté pour surveiller. C'est le principal critère d'arbitrage entre différentes techniques de surveillance (Aiello, 1993 ; Carayon, 1993).
Les informations issues de la surveillance sont utilisées pour réaliser des inférences sur le comportement des individus et elle est la base d'une architecture plus importante qui est celle des incitations : l'utilisation des sanctions ou des récompenses est assujettie à l'interprétation des informations contenues dans le système de surveillance. Cette symbiose entre la surveillance et les incitations permet de garantir l'acceptation de la surveillance dans la relation de travail (Chen et Sanders, 2007) : les règles sont ainsi précisées et font partie intégrante de la récompense des comportements qui respectent les normes de l'organisation, et notamment l'effort fourni au travail, c'est-à-dire la rémunération. Cette conception de l'organisation permet d'aborder la relation de travail comme une relation d'agence.

Dans ce chapitre, le type d'information surveillée par les organisations que nous étudions est la performance des employés. On définit la performance au sens de productivité, et plus particulièrement de productivité apparente du travail. Traditionnellement, on opère une séparation entre plusieurs types de productivité : l'efficience productive, qui mesure la production rapportée à la quantité de facteurs de production utilisée ; l'efficience économique, qui mesure la quantité de production rapportée aux coûts des facteurs de production ; l'efficacité productive, qui mesure le résultat de la production par rapport à l'utilisation des facteurs de production ; l'efficacité économique, qui rapporte le résultat de la production au coût de production (De Bandt, 1988).
Plus généralement, il s'agit de rapporter l'évolution d'un volume d'activité à l'évolution de la quantité de travail exécuté pour produire cette activité et de mesurer la quantité de ressources utilisée par les employés. L'origine de la mesure de la productivité n'est pas précise, mais les premiers enregistrements systématiques de l'activité remonteraient à 1891 aux États-Unis, lorsque le Bureau des statistiques du travail rassembla les premières statistiques détaillées sur la productivité pour mettre en parallèle la production et le nombre d'ouvriers travaillant à cette production dans plusieurs branches d'activités. La productivité est une préoccupation à la fois microéconomique, parce qu'elle participe à concevoir la gestion la plus efficace des ressources pour les producteurs, et macroéconomique, puisqu'elle permet de déterminer la contribution à la croissance de la richesse et à la fixation des prix (Vincent, 1965).
La surveillance et l'évaluation des performances des travailleurs ne sont donc pas nouvelles (Ball, 2010). La poursuite de la profitabilité de l'organisation grâce à une productivité efficace ou efficiente a impliqué une rationalisation de l'activité, c'est-à-dire une organisation scientifique de l'utilisation des ressources destinées à la production. Les récents développements de la surveillance du travail sont en ligne directe avec les travaux de F. Taylor au XXème siècle.

Traditionnellement, cette surveillance se caractérisait, par exemple, par le comptage des heures, la pesée des produits fabriqués, une sonnette qui avertissait de l'ouverture d'une caisse enregistreuse, etc. Elle se caractérisait également par les déplacements des employeurs et des surveillants pour vérifier l'activité des employés, c'est-à-dire pour examiner si le travail était réalisé dans les conditions demandées. C'était une surveillance physique et manuelle, qui reposait essentiellement sur des capacités humaines.

Depuis environ trente ans, la principale évolution des techniques de surveillance de la performance est l'utilisation d'appareils électroniques, notamment l'ordinateur, lequel, lorsqu'il est programmé, permet de traiter plus rapidement l'information numérique. Le développement concomitant de l'informatique, c'est-à-dire la science qui permet la création de ces programmes qui fournissent une séquence d'instructions pour enchaîner des opérations visant à répondre un problème donné, a permis le développement rapide de nouveaux modes de surveillance. La réduction du coût des matériels électroniques, la vitesse d'exécution des programmes informatiques et les changements dans les façons de travailler ont rendu plus faciles l'enregistrement électronique des comportements et la création d'indicateurs de mesure de ces comportements.

Les techniques de surveillance électronique de la performance sont variées (Kizza et Ssanyu, 2004) : utilisation de caméras, enregistrement des conversations téléphoniques, surveillance de l'utilisation d'Internet (consultation de sites, rédaction de messages), vérification du contenu des ordinateurs, capture de mouvement, localisation via la position GPS. Ces caractéristiques changent ainsi profondément les rapports de travail entre les managers et les exécutants. Ce changement de support fait également naître une double interface : l'outil qui sert à réaliser la production permet de rendre compte de cette production.

La raison qui conduit les organisations à surveiller les comportements des employés en général et leurs performances en particulier est que la surveillance permet de procurer une triple protection pour assurer sa survie (Ball, 2010 ; Chen et Sanders, 2007 ; Tomczak et al, 2017 ; West et Bowman, 2014).

La surveillance permet tout d'abord à l'organisation de se prémunir contre les comportements malveillants, c'est-à-dire face à des situations d'espionnage industriel, de vol de secrets industriels ou de sabotage de l'appareil productif.

La surveillance de la performance des employés joue également un rôle de protection des employés contre les accidents liés à la production. Elle procure donc un gain en termes de conditions de travail.

La surveillance permet enfin de vérifier que les employés atteignent les objectifs de l'organisation, au premier rang desquels se situe la profitabilité (Stahl et al., 2005). C'est ce rôle qui nous intéressera dans ce chapitre. La surveillance permet de se prémunir d'un comportement des employés qui ne coïnciderait pas directement avec les attentes de l'organisation, c'est-à-dire d'un risque que ce comportement soit contre-productif et que l'effort fourni par les employés ne soit pas au niveau de celui demandé par l'organisation. Cet aléa moral est lié à l'anticipation de la désutilité à l'effort. Dans l'objectif de profitabilité, la surveillance contribue en premier lieu à inciter à ce que l'effort fourni par l'employé soit positif, à un niveau qui servira à maximiser la rentabilité de l'organisation. À la suite du rôle d'activateur de l'effort, la surveillance joue également un rôle de maintien de l'effort au niveau qui est souhaité, l'idée étant qu'il faut réguler l'effort pour qu'il soit au constamment au même niveau. Elle joue donc un rôle de frein à la diminution de l'effort, en cherchant en permanence le moyen de créer une norme idéale et atteignable. Enfin, la surveillance permet d'augmenter le niveau de performance ainsi que les gains de productivité. La surveillance joue alors un rôle d'évaluation afin de pouvoir dépasser les objectifs qui sont préalablement fixés, c'est-à-dire que lorsqu'un individu connaît son niveau de performance, une norme supérieure lui est demandée pour améliorer ses résultats.

En ce qui concerne l'adoption rapide et étendue de la surveillance électronique, cela s'explique par le changement de son intensité (Alge et al., 2004). Le caractère électronique de la surveillance permet désormais de suivre en temps réel, c'est-à-dire instantanément, de manière continue, une quantité plus importante d'informations sur chaque individu. La surveillance électronique permet ainsi de collecter de l'information « à la finesse du laser et avec une capacité d'absorption d'une éponge » (Marx, 1999). Le changement de support de la surveillance de la performance, en passant ainsi du physique à l'immatériel, la rend moins dépendante des biais cognitifs de la surveillance humaine : elle devient ainsi davantage précise, neutre et objective sur la mesure de la performance des individus. Ce critère renforce l'équité de l'application du système de récompenses et de sanctions, parce qu'elle donne des informations pertinentes sans jugement a priori (Marx, 1999).

Enfin, les critères de facilité de déploiement, la vente de solutions logicielles clés en main, la miniaturisation des outils de surveillance, le caractère ludique et d'accompagnement sont les autres déterminants de l'implémentation des techniques de surveillance de la performance (Tomczak et al., 2017).

2. Un secteur d'activité recourant traditionnellement à la surveillance des performances au travail : les centres d'appels

Il existe un secteur d'activité dans lequel la recherche s'est largement développée pour étudier les effets de la surveillance électronique de la performance (Ball et Margulis, 2011 ; Russel, 2008) : il s'agit des centres d'appels. Apparu aux États-Unis dans les années 1970, et implanté en France à partir des années 1990, ce secteur d'activité a connu une forte expansion en termes de production et d'emploi partout dans le monde.

Un centre d'appels est une structure basée sur le téléphone et l'ordinateur, qui permet une communication directe et à distance entre des clients et un conseiller téléphonique, afin de répondre à un besoin ou de développer la relation clientèle que ce soit par exemple sous la forme de service après-vente, de démarchage de futurs consommateurs des produits des organisations ou de définition du marché. Les secteurs d'activités qui disposent de ces services sont variés : banque et assurance, télécommunication, énergie et santé (Lourel, 2006). L'architecture technique assure de manière automatique, continue et en temps réel la répartition des appels entre les conseillers téléphoniques (Callaghan et Thompson, 2001). Cette activité n'est pas née par hasard, mais dans un contexte social et économique : la délivrance de l'information à coût efficace, dans un environnement dérégulé et où la compétition entre les producteurs s'est accrue (Russel, 2008).

La recherche autour de cet objet « nouveau » a fait naître une dichotomie entre le centre d'appels comme un objet d'étude sur ces modes de gestion et de production et le centre d'appels comme une métaphore de plus vastes mutations sociales, dans lesquelles s'inscrirait systématiquement l'utilisation de la surveillance des performances (Russel, 2008). Trois évolutions socio-économiques majeures qui s'incarnent dans les centres d'appels permettent ainsi de justifier l'utilisation de la surveillance des performances : le développement des dispositifs issu du couplage téléphonie-informatique dans les organisations ; la tertiarisation de l'économie ; le recours à la sous-traitance.

La caractéristique principale des centres d'appels est le couplage téléphonie-informatique (Taylor et Bain, 1999), c'est-à-dire l'association entre le système de transmission sonore à distance de l'information et le traitement automatique et rationnel de l'information. Cette articulation se caractérise par l'utilisation simultanée du téléphone et de l'ordinateur. Cette innovation technologique est au service de la production du centre d'appels : les conseillers qui traitent les appels disposent d'informations préenregistrées dans leurs ordinateurs sur leurs correspondants au téléphone. Cette intégration du téléphone et de l'ordinateur a ainsi permis de concevoir le centre d'appels comme une file d'attente qu'il s'agit d'optimiser pour traiter efficacement les appels émis et reçus en utilisant des modèles de simulation ou de programmation mathématique (voir par exemple le numéro spécial de Management Science n°54, Koole, 2008).

Cependant, l'association du téléphone et de l'ordinateur sert également à surveiller les performances des salariés (Chaulet, 2006). En effet, l'activité des conseillers téléphoniques laisse des « traces » électroniques qui servent ensuite à mesurer des indicateurs de gestion (Ball et Margulis, 2011 ; West et Bowman, 2014 ; Perkins, 2013). Ces indicateurs sont à la fois quantitatifs, lorsqu'il s'agit de mesurer la durée des appels, la volumétrie des appels, le traitement (par exemple, y appels traités en x secondes), le temps pris avant de répondre à un appel, les durées de connexion, les heures de connexion et de déconnexion, et qualitatifs, lorsqu'il s'agit d'écouter les conversations pour analyser le contenu des appels. Les résultats issus des mesures électroniques sont par la suite transmis en retour aux employés pour montrer leurs performances (Wickham et Collins, 2004). C'est donc l'existence même de l'outil technique permettant la surveillance qui contribue à son utilisation.

De plus, les centres d'appels sont une branche d'activité qui rentre dans la catégorie des services. L'output des centres d'appels constitue la prestation d'un service, qu'il soit de résoudre un problème ou de trouver de nouveaux débouchés. Ces deux types d'output correspondent à la définition théorique de l'économie des services (Djellal et Gallouj, 2007) : l'output est flou, c'est-à-dire qu'il est immatériel, non stockable et que c'est un processus instable, un changement d'état ; l'output a des effets dans le temps, c'est-à-dire qu'il produit une conséquence directe à court terme (conclusion d'un accord, satisfaction du client) et indirecte à long terme sur le consommateur (vente véritablement réalisée) ; l'output est coproduit, c'est-à-dire qu'il dépend du consommateur, notamment sur sa capacité à répondre à un appel, ou à formuler une demande précise et il dépend de ses besoins particuliers. Le service rendu par les centres d'appels est spécifique, différencié et personnalisé.
Ainsi, comme tout le secteur tertiaire, les centres d'appels se caractérisent par une difficulté à identifier le numérateur de la productivité. L'enjeu est donc de mesurer ce que réalise chaque individu pour être capable d'identifier qu'elle est la valeur ajoutée de chaque individu (Livian et Baret, 2002). Notons également que le secteur tertiaire s'identifie par une faible et lente augmentation de ses gains en productivité. Cette caractéristique renforce l'idée selon laquelle la mesure de la productivité du service rendu doit servir à légitimer l'activité et sa contribution à la croissance, au partage de la valeur ajoutée et à sa contribution à l'emploi. Il est donc important de pouvoir mesurer ce que réalise concrètement le conseiller téléphonique, ce qui contribue à utiliser un dispositif de surveillance des performances.

Enfin, les centres d'appels se caractérisent par un recours élevé à la sous-traitance de la relation client. La sous-traitance se définit comme étant le fait pour une organisation, dite donneur d'ordres, de faire appel à une organisation extérieure, dite sous-traitante ou preneur d'ordres,

pour exécuter, selon un cahier des charges préétabli et pour une durée déterminée, tout ou partie des actes de production ou de services, soit parce qu'elle ne dispose pas en interne des compétences spécifiques pour réaliser cette activité (sous-traitance de spécialité), soit parce que l'organisation externe est plus compétitive (sous-traitance de capacité) (Algava et Amira, 2011). Le recours à la sous-traitance s'explique notamment par l'instabilité de l'environnement économique, c'est-à-dire une variation inattendue de la demande à laquelle font face les entreprises. Résister à cette instabilité passe par la réduction des coûts de production, et notamment la réduction du coût du travail (Walsh et Deery, 2006). Les coûts de personnels sont représentés par le recrutement, la sélection, la formation et le management des employés ainsi que par le coût des prélèvements sociaux. Or, pour les centres d'appels, les coûts de personnels représenteraient 70 % des coûts totaux (Gans et Zhou, 2005). C'est donc un secteur qui est plus susceptible d'avoir recours à la sous-traitance (Kinnie et al., 2008).

Le taux de recours à la sous-traitance et l'intensité moyenne de ce recours (mesurée, à partir des enquêtes annuelles d'entreprises [EAE], par la part des dépenses de sous-traitance dans le chiffre d'affaires) se sont fortement accrus entre 1984 et 1995 avant de se stabiliser entre 1995 et 2003 (Perraudin, Thèvenot et Valentin, 2013). En ce qui concerne les centres d'appels, la sous-traitance de la relation client est devenue une industrie globale représentant en 2003 un milliard de dollars au niveau mondial (Gans et Zhou, 2005). Elle représenterait 10 à 15 % de l'emploi en centre d'appels et elle augmenterait de 15 à 20 % par an (Walsh et Deery, 2006)[12].

La littérature académique s'est employée à trouver les moyens de piloter efficacement la relation de sous-traitance, notamment à cause des difficultés de coordination entre les deux parties sur la qualité de la prestation de service. Cette relation s'apparente ainsi à une relation d'agence (Chanson, 2003) : l'entreprise donneur d'ordre obtient par un contrat d'externalisation de la société prestataire un engagement sur des objectifs spécifiés. Cependant, le sous-traitant réalise des opérations qui ne sont pas observables par le client : recrutement, formation et investissement dans la technologie. Dans ce cas, le donneur d'ordre utilise le contrat de sous-traitance pour influencer le comportement inobservable. Dans ce contexte, la surveillance des performances des conseillers est une surveillance des indicateurs des contrats de sous-traitance négociés au niveau global. L'organisation opère un transfert de la réalisation des objectifs de contrat de service défini à un niveau global à un niveau individuel. Les employés se trouvent

[12] Dans les arrangements traditionnels d'externalisation de la relation client (Askin, Véricourt et Karaesmen, 2008 ; Hasija, Pinker et Shumsky, 2008), le donneur d'ordres spécifie les objectifs de qualité de service en termes de file d'attente (durée d'attente avant décrochage, durée moyenne d'attente), d'abandon, de traitement des appels (durée moyenne de réponse), de satisfaction des clients, et les pénalités et récompenses associées à chaque indicateur. Un appel pris génère un revenu pour l'opérateur et le coût du service est réparti entre les deux contractants. Le paiement peut s'effectuer selon un critère de volume d'appels traité ou pour la durée des appels traités. Les contrats de sous-traitance de la relation client correspondent ainsi à la fois au critère de sous-traitance de capacité, c'est-à-dire la sous-traitance de la fluctuation et des pics d'activités (*pay-for-capacity contract*), et au critère de spécialité lorsque les appels les moins importants (sous-traitance de la base) sont traités par le preneur d'ordre (*pay-for-job contract*).

tiraillés entre les objectifs de leurs donneurs d'ordre, l'organisation de travail de leur propre employeur et le besoin des clients avec lesquels ils interagissent (Kinnie et al., 2008). Ils sont emboîtés dans une double relation d'agence : la relation d'agence de sous-traitance et la relation d'agence de la relation d'emploi (Gillet, Greenan et Le Gall, 2015).

En définitive, ces trois caractéristiques spécifiques aux centres d'appels entraineraient une intensité élevée de la surveillance électronique des conseillers téléphoniques. Cependant, cette intensité n'est pas sans conséquence sur la qualité de vie au travail des employés.

3. L'approche par la qualité de vie au travail des salariés

La qualité de vie au travail est un concept récent pour lequel il n'existe pas encore de consensus scientifique et qui alimente de nombreux débats académiques (Findlay et al. 2013). Traditionnellement, l'approche par la qualité de vie au travail est une alternative à l'approche scientifique du travail pour incorporer les préférences individuelles dans la conception de systèmes organisationnels en identifiant les aspects du travail qui sont perçus comme désirables (Levine, Taylor et Davis, 1984). Parce que ce concept peut se rapprocher de celui de bien-être dans une organisation productive, il est par essence un concept multidimensionnel. La difficulté d'une définition unique n'empêche pas de proposer des dimensions à la qualité de vie au travail : la rémunération, la sécurité au travail, le développement et le déploiement de capacités individuelles, l'engagement dans le travail, la possibilité d'avoir un contrôle sur les tâches, la capacité d'exprimer des opinions et de participer au processus de décision, la satisfaction à réaliser son travail dans de bonnes conditions physiques et psychiques. La définition contient ainsi à la fois le contenu du travail, en tenant aussi bien compte de l'amplitude horaire que des tâches réalisées, et l'environnement du travail.

Les recherches en sociologie du travail ont pu montrer que les centres d'appels étaient l'épicentre de conditions de travail et d'emploi dégradées : la main-d'œuvre est flexible, leur niveau de qualification est peu élevé, leur rémunération est faible, les possibilités d'évolution professionnelle sont peu diversifiées, les conflits sociaux sont nombreux, la cadence de travail est élevée, le contrôle sur leur activité est peu développé, le stress et la pénibilité dans le travail sont importants. Par conséquent, les sociologues ont pu écrire que les centres d'appels étaient des « usines modernes » (Buscatto, 2002), où le travail était taylorisé : c'est la concrétisation de l'organisation scientifique du travail, où l'autonomie des travailleurs est réduite au traitement de la file d'attente des appelants. Le développement de la sous-traitance renforce cette image de conditions de travail dégradées particulièrement dans les centres d'appels (Batt et al, 2005 ; Connel et Hannif, 2008) : les sous-traitants sont des entreprises plus petites, employant davantage d'employés moins qualifiés que les donneurs d'ordres, où le travail est plus prescrit et contrôlé et

les objectifs moins atteignables, le rythme de travail plus soutenu, plus de risques d'accidents du travail, et l'insécurité socio-économique plus importante. De façon générale, la satisfaction au travail des salariés semble également un peu moins élevée chez les sous-traitants que chez les donneurs d'ordres. Dans les centres d'appels sous-traitants, le turn-over est plus important et la rémunération plus faible que les centres de contact internes : en 2003, le salaire moyen des conseillers en centre d'appels sous-traitant était de 15 520 € contre 17 940 € dans les centres internes (Lanciano-Morandat, 2005).
Cette image est néanmoins à nuancer, parce que « tous les centres d'appels ne se ressemblent pas » (Holman et al., 2007). De plus, le rapport rédigé par Lanciano-Morandat (2005) fait état pour la France, d'un niveau de qualification élevé (Bac+2), d'une sécurité de l'emploi élevé (75 % des salariés en CDI). Cependant, le rapport confirme l'idée que les employés ont peu de prise sur leur travail et que l'autonomie est peu développée.

Dans ce cadre, nous formulons l'hypothèse que l'intensité de la surveillance des performances serait un déterminant de la plus ou moins grande qualité de vie au travail et ce serait dans les centres d'appels sous-traitants que se situeraient les niveaux les plus intenses de surveillance électronique de la performance qui interagiraient négativement avec la qualité de vie au travail des employés. Les désavantages créés par la surveillance électronique des performances seraient par conséquent supérieurs aux avantages pour l'organisation, si cela se traduit par une prise en charge des frais de soins, par de l'absentéisme, et *in fine* par une réduction de la productivité et des performances (Holman et al., 2002).

Par conséquent, le modèle productif adopté par les centres d'appels en général et les centres d'appels sous-traitants en particulier qui utilise de manière intensive la surveillance électronique des performances semble améliorable, parce qu'il fait craindre des interférences avec les conditions de travail des employés. Nous souhaitons ainsi déterminer dans quelle mesure la réduction de l'intensité de la surveillance permanente de la performance individuelle des conseillers téléphoniques modifierait les rapports entre les superviseurs et les salariés afin d'améliorer la qualité de vie au travail sans nuire à la performance globale de l'activité.

Nous avons tâché de répondre à cette question en expérimentant une réduction de l'intensité de la surveillance électronique des performances dans un centre d'appels sous-traitant localisé en Côte d'Or. Cette expérimentation de terrain a tenté d'ouvrir la boîte noire de la surveillance au travail, afin de révéler les coûts cachés de la supervision électronique des performances et pour fournir des pistes de recherche pour modifier le modèle économique des centres d'appels en général et des sous-traitants en particulier.

Le chapitre débute par une revue de la littérature théorique et appliquée qui vise à approfondir les effets de la surveillance des comportements au travail sur la performance et la qualité de vie au travail. Il se poursuit par une présentation du terrain ainsi que du protocole expérimental. Nous présentons ensuite les méthodes d'identification des effets. Les principaux résultats sont détaillés dans une quatrième partie. Enfin, nous interprétons les résultats à la lumière de notre travail d'enquête de terrain et des arguments développés dans la littérature académique.

II. Revue de la littérature

Plusieurs courants de l'économie et de plusieurs autres disciplines, comme la psychologie ou la sociologie, peuvent être utilisés afin de comprendre les effets de la surveillance, à la fois dans le cadre du travail en général, et dans le cadre du travail dans les centres d'appels en particulier. Les théories sont ici aussi bien complémentaires que contradictoires parce qu'elles prévoient des effets positifs ou négatifs sur les individus. Il existe donc une ambiguïté sur les effets de la surveillance.

1. Les effets théoriques ambivalents de la surveillance électronique de la performance

a. L'effet profitable de la surveillance de la performance dans la relation d'emploi

Le principal argument théorique de l'activation d'un dispositif de surveillance au travail est de permettre de vérifier que les employés atteignent les objectifs de production que la firme qui les emploie leur a fixés, et d'examiner dans quelle mesure ils ne sont pas remplis. C'est ce qui est présenté dans le cadre général de la théorie d'agence dans l'entreprise (Alchian et Demsetz, 1972 ; Calvo et Wellisz, 1978 ; Hölmstrom et Milgrom, 1991 ; Jensen et Meckling, 1976).

La relation d'emploi est traditionnellement décrite comme une relation asymétrique. L'employé est dépeint comme un acteur mu par son intérêt individuel et qui exploite toutes les opportunités à sa disposition pour augmenter son bien-être, en arbitrant le temps dont il dispose entre travailler et l'inactivité, le loisir. Dans ce cadre, le travail procure de l'insatisfaction, au contraire du loisir qui offre un gain d'utilité. Au travail, l'employé cherchera à réduire son effort. Ce niveau d'effort n'est pas parfaitement connu par l'employeur, qui cherche à maintenir un certain niveau de productivité afin d'obtenir un niveau de rentabilité maximal de son activité. Dans ce contexte, l'asymétrie d'information entre un agent, l'employé, et un principal, l'employeur, réside sur le niveau d'effort, c'est-à-dire que l'employeur envisage le risque que l'employé ne fournisse pas l'effort demandé. Il existe donc un aléa moral.

Pour réduire cet aléa moral et faire converger les intérêts des deux parties de la relation, le principal a à sa disposition un système d'incitations monétaires, sous la forme d'une rémunération variable dépendante du niveau d'effort par exemple, ou non monétaires, comme le licenciement. Ces incitations reposent sur le dispositif de surveillance et de vérification de l'effort qui agit également sur le comportement des employés à travers deux canaux (Becker, 1968). Ex ante, ce dispositif de surveillance génère un effet de dissuasion qui menace les agents de choisir un comportement opportuniste : le fait de savoir qu'il existe un contrôle n'incite pas à agir dans son propre intérêt. Ex post, lorsque le contrôle détecte des tricheurs rationnels, les sanctions qui sont activées les empêchent de reproduire ce mauvais comportement[13].

In fine, la surveillance envoie un signal qui concourt à ce que les employés qui recherchent leur intérêt personnel poursuivront également les intérêts de leur employeur (Eisenhardt, 1985). Les dispositifs de surveillance constituent donc un paramètre qui rentre dans le contrat optimal de travail.

Au fur et à mesure du progrès technique, les employeurs ont substitué une supervision physique et directe à une supervision indirecte et électronique (Ball, 2010). La supervision électronique réduit les coûts de collecte et de stockage des informations sur le comportement tout en augmentant la quantité et la qualité de ces informations. Dans le cadre du travail, elle est ainsi capable de mesurer en temps réel la performance des individus. Théoriquement, l'alignement des intérêts est plus facilement atteignable, en augmentant l'effet potentiellement disciplinant du contrôle.

b. Les effets nuisibles de la surveillance de la performance au travail

Nous avons décrit dans l'introduction générale dans quelles conditions les dispositifs de vérification de l'information pouvaient provoquer des effets négatifs dans la relation d'agence. La théorie d'éviction de la motivation intrinsèque (Frey, 1993) et la théorie des coûts cachés (Falk et Kosfeld, 2004) de la surveillance s'appliquent directement à la relation de travail, puisque c'est essentiellement cette relation que conceptualisent ces travaux théoriques.

Dans une relation d'emploi, il existe une relation interpersonnelle entre l'employeur et l'employé sous la forme d'un système d'attentes réciproques, ce qui produit un contrat psychologique. Pour les employés, les attentes envers le principal peuvent se concevoir au niveau des conditions de travail offertes par l'employeur (amplitude horaire, espace aménagé pour le travail, rémunération)

[13] En psychologie, la théorie du renforcement et la théorie de la dissuasion (Alge et Hansen, 2014) sont en adéquation avec l'effet disciplinant de la surveillance décrit par les modèles principal-agent. Si le système de surveillance identifie et punit les comportements indésirables, l'occurrence de ces comportements est réduite (effet de renforcement). Si le système de sanctions est communiqué, alors il augmente la sévérité perçue de la détection des comportements déviants et décourage les individus de le faire (effet de dissuasion).

ainsi que d'une attente de gratitude et de reconnaissance dans le travail réalisé. Pour les employeurs, les attentes envers l'agent se situeraient au niveau de la qualité du travail accompli et du niveau d'engagement, sous la forme par exemple d'un développement de la culture d'entreprise et de la solidarité entre les membres de l'organisation. Le renforcement de la surveillance de la part de l'employeur, au-delà d'un niveau socialement partagé et hors du contrat, serait alors perçu négativement par les employés : cela traduit une diminution des marges de manœuvre pour accomplir le travail en autonomie. Par conséquent, l'accroissement de l'intensité de la surveillance de la performance altèrerait le comportement des agents, qui réduiraient leur motivation intrinsèque à réaliser les tâches qui leur sont assignées (Frey et Jengen, 2001).

Selon cette théorie, la motivation intrinsèque repose sur trois besoins fondamentaux : le besoin de compétence, le besoin d'autonomie et le besoin de relation. L'effet d'éviction produit par la surveillance fait référence aux processus psychologiques qui conduisent à une diminution de la motivation intrinsèque : l'auto-évaluation et l'auto-détermination. Le premier effet se produirait lorsque les agents estiment que leur implication et leurs compétences ne sont pas évaluées à leur juste valeur par le principal (Ryan et Deci, 2000 ; Gagné et Deci 2005). Le second effet viendrait du fait que le principal limite l'autonomie des agents en les forçant à se comporter d'une certaine manière sans raison légitime (Schnedler et Vadovic, 2011). Un style de management coercitif, qui implique des procédures strictes et rigides, des évaluations forcées ou des objectifs imposés et des récompenses intangibles, produirait un effet d'éviction sur la motivation intrinsèque par rapport à un soutien de l'autonomie qui renforcerait cette motivation (Arnaud et Chandon, 2013). Une explication complémentaire de la réduction de la motivation intrinsèque réside dans la violation de normes sociales de réciprocité ou d'équité qui affecte le comportement spontané de dons contre dons entre l'agent et le principal (von Siemens, 2013)[14].

De plus, la littérature des coûts cachés de la surveillance suggère que le niveau de surveillance peut être sous-optimal : si une partie du coût de la surveillance est caché pour le principal, alors ce dernier ne le prend pas en compte lors du choix du dispositif de surveillance et cela va vraisemblablement se traduire par un niveau de surveillance au-delà de son niveau d'équilibre (Falk et Kosfeld, 2006). Cette idée se traduit, par exemple, par l'investissement dans une

[14] La réactance psychologique est un autre processus qui intervient dans les réactions comportementales à la surveillance (Brehm, 1966). Les individus, qui ressentent que leur liberté est menacée ou réduite, réagissent en adoptant des conduites contraires à celles qui étaient anticipées. Lorsque les contraintes organisationnelles ne permettent pas de telles réactions, un malaise psychologique est généré, ce qui augmente le niveau de stress et conduit à des stratégies de sortie.

Enfin, selon la théorie de la facilitation sociale (Zajonc, 1965), la surveillance modifierait les réactions des employés en fonction de leur degré de maîtrise sur les tâches à exécuter et des interactions sociales avec le superviseur. Les employés auraient ainsi une réaction de « routine » lorsqu'ils rencontrent des situations familières et des réactions « spontanées » lorsque les situations sont nouvelles. Ainsi, la présence d'une supervision améliorerait l'exécution de tâches simples ou bien maîtrisées par l'employé, mais l'endommagerait lorsque la tâche est complexe, assimilée de façon incomplète ou pendant un processus d'apprentissage.

technique de surveillance de la performance qui engendre une augmentation des coûts de production, plutôt qu'une diminution des coûts liés à une hypothétique occurrence de comportements de débrayage de la part des employés. De surcroît, le simple fait de vouloir surveiller des performances entraine un comportement de surveillance accrue, qui s'autoalimenterait. En définitive, la surveillance réduirait le profit.

Il peut exister deux manières d'expliquer ce phénomène. Bental et al. (2012) expliquent que, lorsque l'intensité de la surveillance du principal n'est pas contractualisée, la relation d'emploi se transforme en un double problème d'aléa moral. Un principal opportuniste augmenterait l'intensité de son contrôle après l'établissement du contrat rémunération-incitation-surveillance pour accroître ses gains. Les chercheurs ont montré que ce cadre de double aléa moral conduit à une sur-surveillance et induit un niveau d'effort sous optimal. La faculté de la surveillance électronique à augmenter l'intensité de la collecte d'information sur les performances serait ainsi un frein à l'effort optimal.

La propension à faire confiance serait un déterminant supplémentaire à une surveillance trop élevée. Généralement, la confiance et la surveillance sont conceptualisées comme fonctionnellement équivalentes : elles sont deux modes alternatifs de gestion de l'incertitude et de la complexité dans les organisations (Reed, 2001). Dans le modèle principal-agent de Falk et Kosfeld (2006), le principal peut choisir entre faire confiance ou surveiller l'agent. Cet arbitrage est guidé par le degré de vulnérabilité du principal aux actions de l'agent et par les anticipations des résultats (Zand, 1972 ; Alge et al., 2004). C'est en suivant cette idée que Alge et Hansen (2014) ont développé un cadre général de cycle confiance-surveillance à l'intérieur duquel il y a un potentiel cercle vicieux de la surveillance et de la méfiance : le surveillant réagirait aux ajustements d'effort des agents en se méfiant de leurs résultats futurs tout en augmentant l'intensité du contrôle[15].

c. L'ambivalence théorique des effets de la surveillance de la performance : une porte ouverte pour la modélisation

D'un point de vue théorique, les effets de la surveillance en général et des effets de la surveillance de la performance en particulier seraient donc ambivalents. Sewell et Barker (2006) ainsi que Sewell et co-auteurs (2012) ont analysé cette ambivalence de la surveillance de la performance en se rapportant à Foucault (1975). Ils définissent la surveillance comme « les moins nombreux regardant les plus nombreux » et identifient deux systèmes alternatifs de représentations dans les organisations : la représentation attentionnée (*caring*) et la représentation coercitive (*coercion*) liées

[15] La recherche en science de l'éducation a également exploré les motivations d'un comportement de sur-évaluation des enseignants, Deci et al. (1982) ont démontré que lorsque les enseignants se sentent découragés par leurs supérieurs hiérarchiques, ils ont tendance à davantage contrôler leurs élèves et particulièrement lorsque la hiérarchie insiste sur la responsabilité qui incombe à leurs élèves à se conformer aux normes de réussite.

aux deux facettes du pouvoir de supervision. Dans la vision attentionnée, l'objectif est de réduire le risque d'aléa moral et de maintenir l'intégrité du projet de l'organisation. Dans la vision coercitive, l'objectif est de forcer les employés à travailler aussi durement qu'ils le peuvent y compris contre leur propre intérêt. Les membres de l'organisation ont donc une expérience paradoxale de la surveillance qui peut agir dans ses deux versants simultanément.

Malgré cette tentative de conciliation des deux versants de la surveillance, le débat qui oppose les théoriciens de l'effet disciplinant de la surveillance et les théoriciens des coûts cachés de la surveillance porte seulement sur l'implémentation de la technologie de surveillance et non véritablement sur leurs interactions avec d'autres variables (Festré et Garrouste, 2015). Il y a donc un champ ouvert pour formuler un cadre théorique plus global qui tiendrait compte de la tension qui existe sur les dispositifs de surveillance.

2. Les preuves empiriques des effets ambigus de la surveillance électronique de la performance

Plusieurs travaux appliqués ont testé ces théories au cours d'expérimentations de laboratoire qui ont des similitudes avec l'environnement de travail des centres d'appels, voire au cours d'expérimentations de terrain in situ. Ces études mettent en lumière les effets positifs de la surveillance sur les efforts des participants aux expérimentations sous plusieurs conditions, tandis que les effets sur le bien-être et la motivation peuvent être le plus souvent néfastes.

a. Les effets de la surveillance électronique de la performance sur l'effort au travail

Dans cette section, nous examinons la recherche empirique qui évalue les impacts de la surveillance électronique de la performance sur l'effort au travail. Nous limitons cette revue aux recherches entreprises dans les centres d'appels et aux expériences de laboratoire qui tendent à imposer un cadre de travail qui présente de nombreuses caractéristiques communes avec le travail dans les centres d'appels. Enfin, nous faisons référence à plusieurs travaux de psychologie, de gestion qui ont également précisé certains modérateurs de la surveillance électronique des performances sur l'effort au travail.

Dans une expérimentation de terrain, Nagin et al. (2002) ont testé l'existence de comportements opportunistes dans les relations de travail de seize centres d'appels d'une entreprise spécialisée dans le démarchage de dons privés. Dans cette entreprise, les conseillers téléphoniques appelaient à partir d'une liste de prospects des donateurs potentiels et leur demandaient une contribution monétaire. L'authenticité des promesses de dons collectées était vérifiée par un rappel téléphonique ou un audit. La fréquence des vérifications à l'aide des rappels n'était pas révélée

aux employés, mais pouvait être déduite par la réduction de la paie à la fin de chaque semaine. Dans quatre sites expérimentaux, la surveillance a été manipulée afin de réduire le taux de vérification. Les résultats de l'expérience montrent que plus le taux de rappel (la surveillance) est faible, plus la proportion de fausses promesses de dons (l'opportunisme) est élevée et que les changements de comportement les plus marqués se produisent à un taux de vérification nul. Ainsi, une surveillance plus intense aurait un impact positif sur la performance. Cependant, le comportement des conseillers des centres d'appels est hétérogène : si certains d'entre eux se comportent comme des tricheurs rationnels, beaucoup ne profitent pas de la réduction de l'intensité de surveillance. De plus, les entretiens complémentaires montrent que les tricheurs rationnels perçoivent plus fréquemment leur employeur comme moins conciliant et plus inéquitable. Ce dernier résultat fait écho au cycle de surveillance-confiance décrit par Alge et Hansen (2014). L'étude insiste sur le fait que la qualité de la relation entre les agents d'appel et les superviseurs est susceptible de modérer l'effet de la surveillance sur les performances.

Dans une expérimentation d'effort réel en laboratoire, Dickinson et Villeval (2008) ont tenté de distinguer l'effet disciplinant de l'effet d'éviction de la motivation intrinsèque engendré par la surveillance. Afin de créer une relation interpersonnelle, les chercheurs ont permis de laisser les « joueurs » avoir des relations sociales, c'est-à-dire qu'ils pouvaient discuter. Ils montrent que la surveillance permet de réduire les comportements opportunistes, mais que les « joueurs » ayant le rôle de principal peuvent choisir un niveau de surveillance plus élevé que celui qui est recommandé pour induire une motivation optimale. En particulier, une distance sociale plus réduite entre un agent et un principal est une condition majeure pour que l'effet d'éviction apparaisse. De plus, ils montrent que l'effet d'éviction fonctionne à travers le mécanisme de réciprocité, parce qu'il apparaît quand les gains du principal augmentent directement et continuellement avec le résultat de l'agent : en effet, dans ces circonstances, les agents peuvent réagir à la méfiance qu'ils ressentent quand le contrôle devient trop intense en réduisant les gains du principal. Ainsi, plus les agents anticipent des gains élevés du principal avec des contrôles plus intenses, moins les agents vont fournir l'effort demandé.

En 2014, D. P. Bhave (2014) a réalisé deux études auprès d'un centre d'appels interne d'une firme financière et d'un centre d'appels de relation commerciale pour examiner dans quelles conditions « la fréquence d'utilisation de la surveillance électronique fonctionnerait comme un "œil invisible" qui alignerait les intérêts intra-organisationnels et améliorerait la performance des employés ». Il a constaté que bien que la surveillance électronique collecte constamment les performances individuelles, l'information doit encore être utilisée pour produire une évaluation générale chaque semaine ou chaque mois pour chaque conseiller téléphonique. Il montre que lorsque les évaluations sont rapprochées, la performance quantitative augmente, c'est-à-dire le nombre

d'appels traités, ou le nombre de demandes résolues. De plus, l'intensité de la surveillance est positivement liée à la qualité des appels des conseillers téléphoniques. Bien que la méthode utilisée ne permette pas de mettre un lien de causalité entre performance et surveillance électronique, ces résultats suggèrent un effet disciplinant de la surveillance.

Enfin, dans une série d'expérimentations de générations de données en laboratoire, des psychologues ont mesuré des effets d'interaction entre la surveillance, le niveau de compétence et la difficulté de la tâche sur les performances individuelles. Les participants ayant un haut niveau de compétence, mesuré par le degré de maîtrise de la tâche, réussissent plus vite sur cette tâche simple lorsqu'ils sont surveillés que lorsqu'ils ne le sont pas (Aiello et Shao, 1993 ; Aiello et Kolb, 1995). Le résultat inverse est également observé pour les participants ayant un faible niveau de compétence, c'est-à-dire lorsque la tâche assignée paraît nouvelle aux yeux des participants (Aiello et Kolb, 1995). Lorsque la tâche est difficile, la surveillance détériore la performance (Aiello et Shao, 1993 ; Douthitt et Aiello, 2001).
Sous une surveillance électronique, la performance serait similaire à celle qui est observée en présence d'un surveillant (Griffith, 1995) et les mêmes interactions entre difficulté de la tâche et surveillance électronique seraient observées (Aiello et Svec, 1993). Cependant, il y a des pics de performance en présence du surveillant, alors qu'elle est plus stable sous une surveillance électronique (Griffith, 1995).

b. Les effets de la surveillance sur le ressenti subjectif de la qualité de vie au travail

Dans cette section, nous présentons les résultats empiriques concernant les impacts de la surveillance électronique de la performance sur la qualité de la vie au travail. Les travaux que nous avons sélectionnés sont principalement issus de la psychologie, de la sociologie et de la gestion.

Dans deux revues de littérature en sciences de gestion, Bates et Holton (1995) et Chalykoff et Kochan (1989) signalent que la surveillance pourrait faciliter une organisation efficace des ressources humaines, lorsqu'elle permet d'aider à détecter les besoins de formation, à développer des retours réguliers (*feedback*) ou à identifier les problèmes que rencontrent les salariés pour réaliser leur travail dans de bonnes conditions, avec des impacts positifs potentiels sur la satisfaction des employés. Cependant, la faculté d'une organisation de surveiller en temps réel les employés permet d'arbitrer vers moins de rémunérations, puisqu'elle n'a plus besoin d'inciter de façon monétaire les employés à agir dans la bonne direction. Chalykoff et Kochan (1989) ont ainsi montré qu'une surveillance étroite de l'activité augmente le turn-over. Par conséquent, la

surveillance renforce l'insécurité dans le travail, le fait de ne pas respecter les indicateurs de performance étant directement sanctionné par l'arrêt du contrat de travail.

Les travaux en psychologie montrent que la surveillance de la performance conduit à un accroissement du stress ressenti par les salariés, c'est-à-dire un état de tension nerveuse dû à une source extérieure agressive, cette source étant l'impression d'un contrôle permanent du travail accompli (Aiello et Kolb, 1995 ; Aiello et Shao, 1993 ; Carayon, 1994 ; Smith et al., 1992). Le stress intervient lorsque la surveillance rend les termes et les conditions du contrat de travail plus difficiles à atteindre et quand la surveillance est inappropriée ou mal conçue. Davidson et Henderson (2000) constatent un impact négatif de la surveillance sur le stress avec des tâches complexes seulement.
Le stress provoquerait également d'autres états de santé émotionnelle et physique plus graves qui sont constatés dans un cadre de travail surveillé et dans les centres d'appels (Chouanière et al., 2011 ; Grebner et al., 2013 ; Holman, 2004) : la fatigue, l'anxiété, l'énervement, la dépression, la peur et des troubles musculosquelettiques.

De plus, la surveillance entre en conflit avec la réalisation du travail. Grosjean et Ribert-Van De Weerdt (2005) précisent que la surveillance est perçue par les conseillers téléphoniques comme plus exigeante que les échanges avec leurs correspondants. La surveillance compromet l'autonomie et les marges de manœuvre dans le travail (Grebner, 2003 ; Perkins, 2013), c'est-à-dire le contrôle individuel sur son activité personnelle, ce qui se traduit par du burn-out, une situation où l'individu est dépassé par son environnement et où la charge de travail augmente.
La surveillance réduit la créativité et l'engagement dans le travail. Arnaud et Chandon (2013) montrent que la surveillance électronique des performances a un effet d'éviction sur la motivation intrinsèque des conseillers téléphoniques.
La surveillance diminue la communication entre les parties de l'organisation, ce qui se traduit par une perte de confiance dans les relations de travail, engendrant des comportements de suspicion permanents sur l'existence d'attitudes opportunistes (Kizza et Ssanyu, 2004) ou un sentiment d'isolement (Aiello et Svec, 1993).

Enfin, la surveillance électronique des performances est potentiellement créatrice de formes nouvelles de conditions de travail dégradées (Jeske et Santuzzi, 2015). La capacité de pouvoir surveiller davantage d'éléments dans le travail fait craindre de collecter des informations supplémentaires sur les individus, sans leur consentement, et cela nuirait à la séparation entre vie professionnelle et vie privée (Chen et Sanders, 2005 ; Stanton, 2000).

La surveillance électronique des performances serait donc l'un des principaux facteurs d'insatisfaction au travail (Alge et Hansen, 2014 ; Chalykoff et Kochan, 1989 ; Grant et Higgins, 1989). Pour Holman (2004), c'est principalement la perception du contrôle qui aurait un effet néfaste et lorsque l'intensité de la surveillance est perçue comme élevée, elle est négativement liée au bien-être.

3. Synthèse

La revue de la littérature théorique et appliquée suggère ainsi que les systèmes de surveillance électronique de la performance sont une « épée à double tranchant ». Leur utilisation peut être contre-productive en générant l'effet opposé à celui qui était souhaité à l'origine de sa mise en œuvre. Les variations d'intensité de la surveillance modifient les réactions comportementales des individus surveillés. La revue de la littérature suggère dans l'ensemble des effets mitigés des dispositifs de vérification électronique de l'information sur les indicateurs de performance individuelle, tandis que la qualité de vie au travail est la plupart du temps faible, voire dégradée. La surveillance au travail qu'elle soit électronique ou physique a donc des effets ambivalents, sur lesquels il s'agit d'arbitrer afin de concilier tous les aspects de l'environnement du travail.

III. Développement des hypothèses

Les mesures de la performance des conseillers téléphoniques dans les centres d'appels sous-traitants sont déterminées contractuellement par les firmes clientes et le centre d'appel et surveillées par le donneur d'ordre : en fonction des contrats de sous-traitance, il existe des sanctions (pénalités de retard, rupture du contrat) si les objectifs ne sont pas remplis. Dans la pratique, l'architecture technique du centre d'appel permet le transfert direct des contraintes du contrat de service à la relation d'emploi : les conseillers téléphoniques sont directement surveillés et évalués sur leurs performances à un niveau individuel, c'est-à-dire que chaque employé doit remplir personnellement des objectifs fixés à un niveau contractuel global. Nous avons développé l'idée que dans les centres d'appels sous-traitants, la relation d'emploi est emboîtée à l'intérieur d'une relation de service avec des donneurs d'ordre.

Or, Frey (1993) et Dickinson et Villeval (2008) ont notamment démontré que les relations interpersonnelles et impersonnelles ne peuvent pas être surveillées de la même manière : le contrôle dans une relation impersonnelle est susceptible d'être à un niveau supérieur que pour une relation interpersonnelle. Dans ce cadre, l'intensité de la surveillance électronique peut être paramétrée à un niveau excessif dans l'environnement de travail des centres d'appels en général et dans les plateformes téléphoniques sous-traitantes en particulier. Selon notre revue de littérature, cette intensité, si elle est paramétrée à un niveau excessif, peut interagir tant à la hausse qu'à la

baisse sur les performances et sur la qualité de vie au travail des conseillers téléphoniques. Ces arguments permettent de formuler les hypothèses théoriques suivantes :

Hypothèse 1 : Dans les centres d'appels sous-traitants, la surveillance électronique est susceptible d'être excessive.

Il est difficile de directement tester cette hypothèse empiriquement. Mais nous en avons extrait deux sous hypothèses qui sont testables. Si le niveau de surveillance est au-dessus de son niveau d'équilibre, alors il entrave la performance parce qu'il génère des coûts évitables. Dans ces circonstances, nous nous attendons à ce qu'une réduction de l'intensité de la surveillance ne réduise pas l'effort des conseillers téléphoniques, mais qu'elle améliore la qualité de vie au travail.

Hypothèse 2 : La réduction de l'intensité de la surveillance électronique dans un centre d'appels sous-traitant ne détériore pas la performance des conseillers téléphoniques.

Hypothèse 3 : La réduction de l'intensité de la surveillance électronique dans un centre d'appels sous-traitant améliore la qualité de vie au travail.

Pour tester ces hypothèses, nous avons mis en place une expérimentation de terrain dans un centre d'appels sous-traitants.

IV. Terrain et méthodologie

1. Contexte de l'expérimentation

a. Une demande originale

Nous avons mis en place une expérimentation contrôlée de terrain par assignation aléatoire (Levitt et List, 2011) dans une plateforme téléphonique, localisée à Dijon en Côte d'Or, à la suite de la sollicitation du dirigeant d'une grande entreprise de la relation client. Fondée il y a 25 ans, l'activité principale de cette entreprise est de permettre à d'autres organisations d'externaliser leur service de relation clientèle. L'organisation a également conçu et commercialisé des logiciels de gestion informatique intégrée de la relation client, ce qui lui assure une position de conseil aux entreprises qui développent leur propre service de contact client en interne. Enfin, l'entreprise s'est investie dans l'amélioration de la qualité de vie au travail des conseillers téléphoniques en créant une ligne de mobiliers spécifiques aux centres d'appels, en développant une politique de formation longue des employés, et proposant dans la mesure du possible des contrats à durée déterminée (fin 2012, 80 % des salariés étaient en CDI sur 500 employés répartis sur cinq

plateformes d'appels). L'entreprise est ainsi à l'opposé de l'image traditionnelle des centres d'appels en France et à l'étranger (Chouanière et al., 2011 ; Cousin, 2002).

Conscient de sa position inédite dans son secteur d'activité, le dirigeant de l'entreprise souhaitait entreprendre une démarche de réflexion afin d'approfondir ce modèle en rupture avec la forme classique des centres d'appels, tout en préservant la performance économique, voire créer un supplément de valeur tout en améliorant la qualité de vie au travail et la qualité de la prestation rendue. Il a donc ouvert les portes de son sixième centre d'appels, terrain initial de nombreuses caractéristiques innovantes, afin de conduire une innovation sociale.

b. La plateforme d'appels : un lieu innovant

Le centre d'appels où a eu lieu l'expérimentation est bâti sur un ancien site industriel, et il est constitué de deux plateaux téléphoniques. Si l'on se réfère à la revue internationale des centres d'appels conduit par Holman et al. (2007) cette plateforme d'appels se distingue du modèle traditionnel de ce secteur d'activité. Ouvert en 2012, le site se caractérise d'abord par des innovations architecturales : des matériaux antibruit couvrent les murs, les sols et les plafonds ; des stations de travail en bois « pour le côté chaleureux » sont organisées en « ilots » ou en « étoiles » dans les plateaux d'appels en *open space* et elles sont conçues pour être ergonomiques et plaisantes, notamment grâce à la personnalisation des postes ; des espaces de convivialité sont conçus pour permettre aux employés de prendre des temps de pause à l'extérieur des plateaux ; des tablettes tactiles sont utilisées à la place des ordinateurs fixes pour certaines campagnes d'appels afin d'apporter de la diversité dans la manière de travailler.

Il se distingue également du modèle traditionnel par sa politique managériale et la gestion de ses ressources humaines : les conseillers téléphoniques sont formés pendant quatre mois et demi en alternance, d'une part par la Chambre de Commerce et d'Industrie (CCI) sur la base du métier de conseiller et d'autre part par des consultants qui viennent au sein de la plateforme d'appels. Les offres de formation continue sont spécifiques à ce centre d'appels et elles sont tournées vers le développement de multicompétences, comme le traitement de plusieurs campagnes d'appels en même temps, et la diversification du type d'appels traités, émis ou reçus.

Enfin, une enquête interne visant à évaluer la pénibilité au travail dressait un portait plutôt positif de centre d'appels comparé aux autres établissements de l'entreprise (Gillet, Greenan, Le Gall, 2015) : en contrôlant des différences d'ancienneté et d'âge entre les différents sites, les conseillers indiquaient un sentiment de travail effectué plus intéressant, des postes de travail plus adaptés, plus de liberté dans l'organisation du travail et la gestion des pauses, une plus grande capacité d'appropriation personnelle des scripts de communication, de meilleures relations avec les collègues, mais une moins bonne reconnaissance du travail par leur hiérarchie et des douleurs ressenties comme plus gênantes dans leur travail. Il est cependant difficile de dissocier l'effet des

innovations organisationnelles de celui de la faible ancienneté des conseillers sur ces résultats positifs concernant la qualité de vie au travail.

c. Le cœur du système inchangé : la surveillance électronique

Néanmoins, les innovations dans l'environnement de travail n'ont pas modifié le cœur de l'organisation et le système de production du centre d'appels, notamment parce que la technologie d'intégration téléphone-ordinateur demeure au centre de l'activité. Il existe ainsi un logiciel de traitement du flux d'appels qui est divisé en deux applications « conseillers » et « superviseurs ». La dimension technique du système de surveillance est complètement maîtrisée par l'organisation grâce au développement en interne de l'outil de traitement des appels ainsi que par le recrutement de deux informaticiens à temps plein. Cet outil dédié initialement au traitement des appels est utilisé pour surveiller électroniquement l'effort des conseillers téléphoniques, parce qu'il permet de mesurer tous les temps passés par les employés dans différents états du système comme un chronomètre permanent et qu'il engendre des traces de l'activité qui sont utilisées pour créer des indicateurs de gestion. Ces indicateurs dérivent des contraintes imposées par les contrats passés entre le centre d'appels et ses donneurs d'ordres. La manière dont les traces de l'activité sont enregistrées dans le système informatique et restituées, en temps réel, aux superviseurs et aux conseillers structure la réalisation des tâches, le système d'attentes et les rapports sociaux.

Le travail des conseillers et l'enregistrement de leur activité dans le logiciel

Les conseillers téléphoniques travaillent ainsi sur leur propre poste de travail avec un casque-micro et ils visualisent sur leur écran les fiches d'appels sur le logiciel de traitement des appels. Ces dernières sont consacrées à communiquer aux conseillers plusieurs informations en fonction du type d'appels. Pour les appels sortants émis depuis la plateforme téléphonique, c'est-à-dire pour vendre un produit par exemple, le conseiller connaît des informations sur l'appelé/client potentiel (nom, adresse, etc.) et la fiche d'appels contient le script de la communication à suivre. Pendant ces campagnes d'appels, ce sont les conseillers téléphoniques qui génèrent les appels à partir d'une liste de prospects, c'est-à-dire de clients potentiels, grâce à deux modes d'émission : le mode manuel permet aux conseillers de composer le numéro du client ; le mode prédictif permet au système de composer automatiquement le numéro de téléphone. L'engagement des conseillers téléphoniques dans le processus de production est un élément central de l'activité dans ce type de campagnes. Pour les appels entrants réceptionnés par la plateforme, c'est-à-dire pour la résolution d'un problème, le conseiller connaît également des informations sur l'appelant/client, ainsi qu'un script pour résoudre ce problème. Pendant ces campagnes d'appels entrants, le flux d'appels est géré automatiquement : les appels sont assignés aux conseillers en fonction de leurs états dans le

système. C'est un processus de production sur lequel les conseilleurs ont peu de contrôle puisque d'une part ils passent beaucoup de temps à attendre un appel pendant les périodes de creux et d'autre part ils peuvent enchaîner les appels sans interruption pendant les périodes de pics d'appels. Le schéma 1 retrace les différents états du système tels qu'ils sont codifiés par le centre d'appels et tels qu'ils sont transcrits dans les applications informatiques qui suivent la production globale en mesurant le temps passé conseiller dans ces états.

Schéma 1 : La ventilation du temps de travail des conseillers dans les états du système

<table>
<tr><th colspan="10">TEMPS DE CONNEXION</th></tr>
<tr><td colspan="3" rowspan="2">Hors Production</td><td colspan="7">Production</td></tr>
<tr><td colspan="3">Occupation</td><td colspan="4" rowspan="2">Non Occupation</td></tr>
<tr><td colspan="2">Travail</td><td>Attente</td></tr>
<tr><td>Pause</td><td>Brief</td><td>Formation</td><td>En Ligne</td><td>En Post Appel</td><td>En Attente</td><td>Demande d'info.</td><td>Temps entre com.</td><td>Consultation des statistiques</td><td>Autres</td></tr>
</table>

Une partie des indicateurs de performance utilisés pour le suivi de la production est constituée de ratios construits à partir de cette ventilation des temps. Par exemple la part de temps passé en post-appel rapporte le temps passé en post-appel au temps passé en production. Ainsi, sur les ordinateurs des conseillers un onglet « statistiques » fournit la ventilation des temps passés par le conseiller depuis qu'il s'est connecté. Par exemple, on y trouve la durée cumulée du temps passé dans chaque état de non occupation et le nombre de passages ainsi que le nombre de communications effectuées, le cumul des durées de communication et la durée moyenne de communication. Pour les appels entrants, le nombre d'appels en attente s'affiche dans un bandeau en bas de l'écran et un code couleur indique quand la file d'attente dépasse un seuil critique.

Le travail des superviseurs et la visualisation de l'activité

Seuls ou en binôme, les superviseurs ont la responsabilité de la gestion des campagnes d'appels, mais pas des conseillers téléphoniques directement, c'est-à-dire qu'en fonction des emplois du temps, la répartition entre les superviseurs et les conseillers téléphoniques s'effectue par le biais des campagnes d'appels, mais il n'existe pas une répartition directe entre les superviseurs et les conseillers. Autrement dit, un conseiller peut être évalué par plusieurs superviseurs pendant la durée d'une même campagne téléphonique.

Sur l'écran du superviseur, l'information mise en scène par l'application de pilotage décrit le travail de tous les conseillers actifs sur les campagnes qui lui sont attribuées (schéma 2). L'architecture de l'application est composée de fenêtres (groupe de téléphonie 1, groupe de téléphonie 2) qui permettent le suivi de chaque campagne téléphonique et des conseillers qui travaillent sur ces campagnes. Chaque ligne d'informations dans ces fenêtres représente le travail d'un conseiller, que le superviseur localise grâce à un identifiant et à son nom. Le superviseur

peut déterminer ainsi dans quel état du système de production se situent un conseiller donné et le cumul de son temps dans cet état. Une colonne « pause », qui mesure le temps de pause cumulé d'un conseiller depuis sa première connexion complète chaque fenêtre de groupe de téléphonie.

Schéma 2 : Application de pilotage qui équipe le poste de superviseur

GROUPE DE TÉLÉPHONIE 1 (ENTRANT)					GROUPE DE TÉLÉPHONIE 2 (SORTANT)				
ID	Nom	État	Durée	Pause	ID	Nom	État	Durée	Pause
X1	X1	En communication	03 :47	01 : 00	X2	X2	En	01 : 00	06 :31
X3	X3	En communication	01 : 15	03 : 00	X4	X4	communication	00 :18	10 :19
X5	X5	Pause	02 :28	09 : 00	X6	X6	Demande info. En post appel	02 :56	5 :57

FILE D'ATTENTE	STATISTIQUES GLOBALES		JOURNAL DES ACTIONS
GT1 :2 post appel, 1 en ligne, 2 attente	Groupe de téléphonie 1		X1 : connexion –
GT3 : 1 en ligne, 0 attente	Nombre d'appels présentés	237	X2 : connexion — en communication —…
GT4 : 3 en ligne, 4 attente	Nombre d'appels pris	192	X3 : connexion — en pause —… – en communication —
	Effectif en pause	1	X4 : connexion —
	Durée d'attente moyenne	4 :30	X5 : connexion —… – en pause —
	Durée moyenne de communication	3 :05	X6 : connexion –
	Durée moyenne de traitement	3 :20	X7 : connexion — déconnexion

L'application de pilotage du superviseur permet aussi de visualiser des indicateurs agrégés au niveau de chaque groupe de téléphonie. Ainsi, dans la fenêtre « statistiques globales », le superviseur accède directement aux statistiques agrégées construites à partir de la ventilation des temps et ceci pour les groupes de téléphonie de son choix. Ces statistiques sont plus pertinentes pour évaluer l'activité sur les appels entrants que sur les appels sortants. Sur les appels sortants, l'information cruciale concerne en effet les issues d'appels qui sont codées à partir du logiciel qui gère les actions et écritures associées à chaque contact. L'une de ces issues est par exemple « contact argumenté » lorsque le conseiller a pu échanger avec l'appelé au sujet de sa campagne, qui peut se conclure par un accord ou par un refus. Ces informations ne sont pas directement accessibles sur l'écran de l'application de pilotage des superviseurs. Pour y avoir accès, ils doivent se connecter à une autre application. Dans la fenêtre file d'attente, le superviseur visualise la situation globale de la file d'attente concernant chaque campagne d'appels entrants. Un code couleur, qui peut être accompagné d'un effet sonore, indique au superviseur si la situation est critique sur certaines de ces campagnes et nécessite une réallocation des ressources. Le superviseur a accès, dans un onglet séparé qui peut ouvrir une autre fenêtre, au journal des actions, qui répertorie, pour chaque conseiller, l'inventaire minuté de la séquence d'actions qu'il a réalisée depuis qu'il s'est connecté au système.

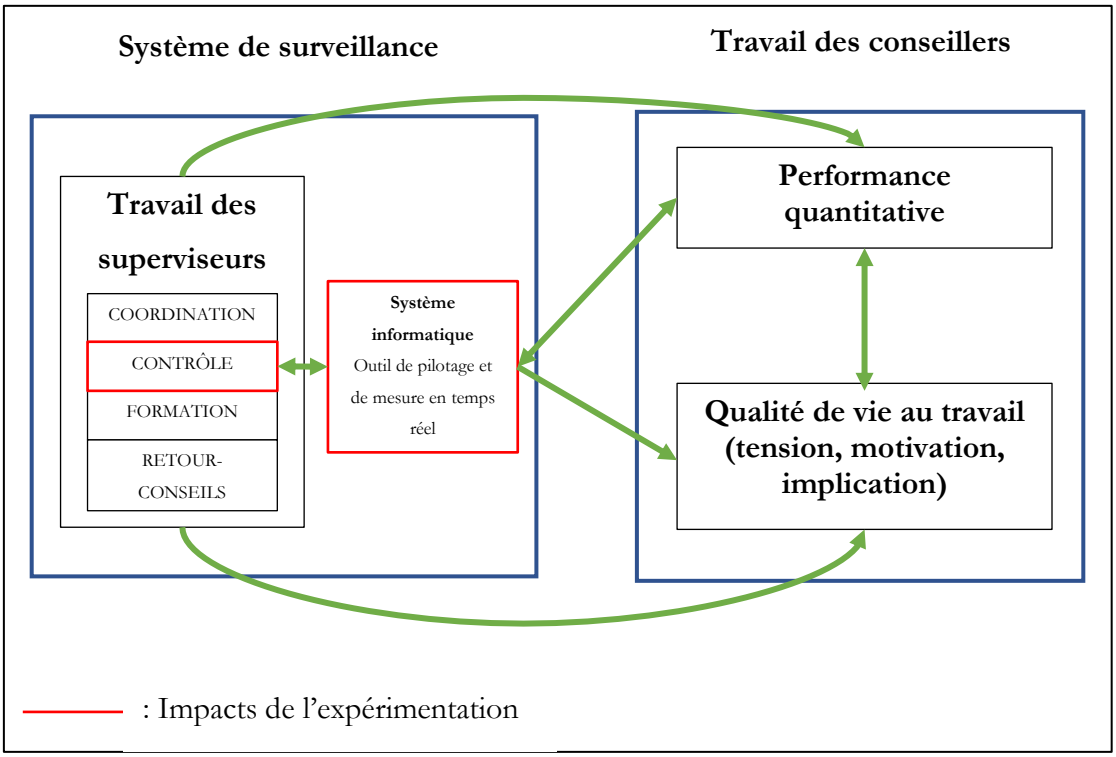
Lorsque survient un problème sur la plateforme d'appels, comme un pic d'appels liés à un incident chez un client ou un fichier d'adresses avec des erreurs d'enregistrement, ce sont les superviseurs qui interviennent sur le plateau. Comme les conseillers sont alloués aux campagnes

selon un volume d'activité planifié, il est souvent nécessaire d'effectuer des réajustements en fonction de difficultés imprévues. Il s'agit là pour le superviseur d'une tâche de coordination.
Les superviseurs remplissent enfin d'autres rôles sur le plateau qui génèrent des tâches qui peuvent elles aussi s'appuyer sur le système informatique : activité de coordination, de *feedback* (brief) et de formation. Ces tâches de soutien permettent de recommander des améliorations dans le travail des conseillers.

La surveillance électronique de la performance

La surveillance électronique de l'activité en général et celle de la performance des conseillers téléphoniques en particulier demeurent centrales dans l'organisation. Nous définissons le système de surveillance électronique de la performance comme les relations entre les superviseurs et les conseillers téléphoniques qui sont modelées par le logiciel de pilotage de l'activité des superviseurs et par l'application de traitement des appels des conseillers téléphoniques qui collectent, enregistrent, traitent, rendent compte et partagent la performance des conseillers. Cette définition compréhensive est proche de celles proposées par Arnaud et Chandon (2013) et par Alge et Hansen (2014).
Le schéma 3 formalise le système de surveillance et son articulation au travail des conseillers. Le système de surveillance est composé des deux éléments centraux : le système informatique d'un côté et le travail de contrôle de l'activité des conseillers réalisé par les superviseurs de l'autre. Les tâches de contrôle prennent appui sur les outils de pilotage et de mesure en temps réel de l'activité que nous venons de décrire.

Schéma 3 : Système de surveillance, travail des conseillers, les modifications de la condition expérimentale



Pour rompre avec ce modèle, il faudrait parvenir à transformer la manière dont le pilotage s'appuie sur les ressources digitales avec l'idée de favoriser la création de valeur plutôt que la surveillance individuelle. La manière dont l'informatique et les superviseurs sont mobilisés pour surveiller les conseillers semble être un verrou à l'évolution en profondeur du modèle organisationnel. Comme la technologie qui en est le vecteur est devenue invisible, c'est un domaine de l'organisation du travail qui, en dépit d'un discours managérial volontariste, demeure d'une grande inertie. La condition expérimentale vise à réduire l'intensité de la surveillance en temps réel du travail des conseillers en limitant les informations en temps réel qui sont mises en scène par les applications informatiques. Nous avons été attentifs à ce que les tâches remplies par les superviseurs, notamment les tâches de coordination, ne soient pas perturbées par les changements mis en œuvre. Dans la condition contrôle, qui sert de point de comparaison dans l'expérimentation, la mise en scène de l'information sur les écrans des conseillers et des superviseurs reste inchangée.

2. Protocole expérimental

Nous avons donc eu l'opportunité de conduire une expérimentation au sein d'une organisation pour tester un changement organisationnel dans un contexte de travail réel. Nous sommes ainsi

des « *insider economists* », le nom prêté aux chercheurs par Kathryn Shaw (2009), qui tentent d'éclairer la boîte noire de la firme. C'est une étape importante dans le champ de l'économie appliquée que de pouvoir tester des hypothèses de réactions comportementales au cœur même d'une organisation productive.

L'expérimentation a d'abord été présentée aux employés et aux dirigeants du centre d'appels comme un moyen de faire émerger un apprentissage collectif sur les manières d'améliorer l'organisation du travail. Dans un premier temps, le protocole expérimental a été décrit au siège social de l'entreprise. Puis les conseillers téléphoniques ainsi que les superviseurs ont reçu une information sur l'expérimentation dans des termes très généraux. Avant le début de l'expérimentation, nous étions présents sur le terrain pour répondre aux questions. Le protocole retenu vise deux objectifs principaux. Le premier est de faire en sorte que les indicateurs quantitatifs demandés par le donneur d'ordres, qui encadrent le contrat de service, ne soient plus utilisés par le superviseur dans un système de surveillance individuelle permanente, parce que cela altère la confiance nécessaire au bon fonctionnement de la relation d'emploi. Le deuxième objectif est de libérer une partie du temps de travail des superviseurs qui serait consacrée aux autres rôles et fonctions qu'ils remplissent sur le plateau.

a. La procédure de changement organisationnelle pour diminuer l'intensité de la surveillance électronique des performances

L'expérimentation a duré 112 jours, du lundi 27 janvier 2014 au samedi 17 mai 2014, durant lesquels nous avons expérimenté une réduction de l'intensité de la surveillance électronique pendant huit séquences de deux semaines consécutives. Afin de réduire l'intensité de la surveillance électronique dans la condition expérimentale, nous avons modifié les paramètres des applications des conseillers téléphoniques et des superviseurs en supprimant l'accès aux informations individuelles en temps réel. La condition témoin préserve l'accès à ces informations individuelles. De plus, nous avons adopté un principe de symétrie : dans les deux applications, nous avons masqué le même type d'information.

En ce qui concerne les paramètres de l'application de traitement des flux d'appels des conseillers téléphoniques, nous avons effectué trois changements sous la condition expérimentale.

En premier, une fenêtre « pop-up » ou « jaillissante » qui indique que « le temps de traitement post-appel est trop long » a été bloquée : en temps normal, elle s'active automatiquement y compris lorsque le seuil temporel n'est pas appliqué et l'inconvénient de cette fenêtre est qu'elle interrompt le traitement des communications. Le fait de la masquer doit ainsi permettre d'avoir davantage d'autonomie dans la conclusion d'une communication.

En deuxième, l'accès à la fenêtre des statistiques en temps réel comme le temps cumulé dans chaque état du système, le nombre d'appels à l'heure, ou la durée moyenne de communication a été supprimé. Ces informations étaient aussi bien bénéfiques parce qu'elles permettent de matérialiser le travail accompli et de réguler l'effort au cours de la journée que préjudiciables parce qu'elles contribuent à contraindre les individus, ce qui est susceptible de provoquer de la tension au travail.
Enfin, le nombre clignotant d'appels en attente qui matérialise la file d'attente pour les campagnes d'appels entrants avec une forte volumétrie d'appels a été supprimé : en effet, le seul moyen de réduire la file d'attente est d'accélérer le rythme auquel ils répondent au téléphone. Plutôt que faire peser l'effort de réduction sur les seuls conseillers, les superviseurs doivent ainsi intégrer leur rôle d'allocation des ressources sur le plateau téléphonique.

En ce qui concerne les paramètres de l'application de pilotage des campagnes d'appels des superviseurs, nous avons effectué quatre changements sous la condition expérimentale.
Premièrement, le chronométrage des temps des conseillers téléphoniques passés dans chaque état du système d'information a été masqué : les superviseurs n'ont plus ainsi plus de signaux visuels pour déterminer les états dans lesquels sont les conseillers de la condition expérimentale.
Deuxièmement, les temps cumulés passés en pause par chaque agent sont également supprimés, sauf pour les campagnes d'appels entrants : en effet, pour ce type de campagne d'appels, les superviseurs ont besoin de cette information afin de distribuer les temps de pause équitablement sans interrompre l'activité durant les pics d'appels.
Troisièmement, la fenêtre affichant les indicateurs de performance individuelle a été masquée : dans ce cadre, nous appliquons notre principe de symétrie, pour que les informations individuelles ne soient plus systématiquement évaluées : comme nous l'avons évoqué précédemment, les indicateurs de performance sont contractualisés à un niveau global et non à un niveau individuel.
Enfin, le journal des actions a été modifié : les mentions « connexion » et « déconnexion » étaient les seules visibles, permettant de vérifier les départs et les sorties des conseillers téléphoniques. Le schéma 4 montre comment la mise en œuvre de la condition expérimentale modifie la mise en scène de l'information de pilotage présentée dans le schéma III.2. Le superviseur suit en particulier deux groupes de téléphonie, un en entrant et un autre en sortant. Il suit le travail de six conseillers (X1, X2, X3, X4, X5, X6) et trois d'entre eux sont dans la condition expérimentale (X1, X4, X6). Pour X1, qui travaille en appels entrants, l'expérimentation masque la durée dans son état de la téléphonie, mais pas son temps de pause cumulé. Dans le journal des actions, seule l'action « connexion » est affichée. Pour X4 et X6, qui travaillent en appels sortants, leurs durées dans les états de la téléphonie sont masquées, ainsi que leurs durées cumulées en pause. Notons

que dans le « journal des actions », un septième conseiller a travaillé cette journée, mais son journal ne comporte que les actions « connexion » et « déconnexion ». Cet individu était dans la condition expérimentale.

Schéma 4 : Application de pilotage qui équipe le poste de superviseur dans la condition expérimentale

Groupe de Téléphonie 1 (ENTRANT)					Groupe de Téléphonie 2 (SORTANT)				
ID	Nom	État	Durée	Pause	ID	Nom	État	Durée	Pause
X1	X1	En communication		01 : 00	X2	X2	En	01 : 00	06 :31
X3	X3	En communication	01 : 15	03 : 00	X4	X4	communication		
X5	X5	Pause	02 :28	09 : 00	X6	X6	Demande info.		
							En post appel		

FILE D'ATTENTE	STATISTIQUES GLOBALES		JOURNAL DES ACTIONS
GT1 :2 post appel, 1 en ligne, 2 attente	Groupe de téléphonie 1		X1 : connexion –
GT3 : 1 en ligne, 0 attente	Nombre d'appels présentés	237	X2 : connexion — en communication —… en post appel
GT4 : 3 en ligne, 4 attente	Nombre d'appels pris	192	X3 : connexion — en pause —… – en communication —
	Effectif en pause	1	X4 : connexion –
	Durée d'attente moyenne	4 :30	X5 : connexion —… – en pause —
	Durée moyenne de communication	3 :05	X6 : connexion –
	Durée moyenne de traitement	3 :20	X7 : connexion — déconnexion

Pour conclure, les superviseurs avaient toujours accès à des informations journalières, hebdomadaires et mensuelles sur les indicateurs de performance individuelle. Les conseillers avaient également un accès à leur comptage de résultat client, c'est-à-dire le nombre d'appels résolus (en entrant) et le nombre de contrats conclus (en sortant), ainsi qu'à leurs performances hebdomadaires et mensuelles. Le protocole ne propose pas une nouvelle façon de travailler, mais il suppose que le temps gagné par les superviseurs en surveillant moins les conseillers peut être réalloué à d'autres tâches du pilotage tout aussi productives : la coordination, la formation ou le *feedback*.

b. Procédure d'assignation aléatoire du changement organisationnel

Nous avons utilisé une procédure d'affectation aléatoire des conseillers téléphoniques entre la condition expérimentale, qui modifiait le paramétrage de la surveillance et la condition de témoin, dans laquelle les outils de mesure de la performance étaient inchangés. L'ingénieur informatique présent sur le plateau téléphonique recevait la liste des conseillers constituant les deux groupes et configurait les applications suivant les changements contenus dans le protocole expérimental. Dans notre étude, les deux groupes agissent dans le même environnement : un lieu de travail unique, des conditions de travail identiques et une quantité équivalente d'interactions avec les chercheurs. Nous considérons donc que cette procédure nous permet d'attribuer un lien de

causalité entre les changements qui sont opérés et les effets que nous mesurons (Rubin, 1974 ; Imbens et Wooldridge, 2009 ; Shaw, 2009 ; Bandiera et al., 2011).

Zaugg et al. (2014) suggèrent que des tirages aléatoires avec remise peuvent être utilisés lorsque l'intervention est fournie séquentiellement et que les participants appartiennent successivement au groupe témoin et au groupe test. Cette méthode est connue sous le nom de « *stepped-wedge* » dans les essais cliniques (Hemming et al., 2014). Afin de déterminer la durée requise pour l'apparition de changements de comportement, nous nous sommes référés aux informations fournies par l'entreprise sur les temps et les durées de formation pour maîtriser de nouvelles compétences. La durée des formations varie entre une demi-journée et quatre jours et une moyenne de quinze jours est nécessaire pour maîtriser pleinement les nouvelles compétences.

Ainsi, à chaque début d'une séquence d'expérimentation de quinze jours, nous avons assigné aléatoirement les conseillers téléphoniques dans les deux groupes, le groupe test et le groupe témoin. Pendant la période expérimentale, tous les participants étaient sous l'une ou l'autre condition et à chaque début de quinzaine les participants étaient réaffectés à l'un des groupes. Dans ce contexte, les conseillers pouvaient ainsi être affectés plusieurs fois à l'une ou l'autre des conditions.

c. Dispositif de collecte de données

Sur la performance des conseillers téléphoniques

L'entreprise nous a accordé l'accès aux données de production journalière qui mesurent la performance. Nous avons conçu avec l'aide de l'ingénieur informatique de la plateforme téléphonique la liste des formats de fichiers informatiques structurant la base de données dont nous avions besoin et nous recevions une extraction du système d'information à chaque fin de séquence de quinze jours. Ces données reflétaient pour chaque conseiller téléphonique travaillant sur une campagne d'appel donnée la ventilation journalière de ses temps de travail pour chaque état du système d'information ainsi que les résultats de chaque contact déclarés par les conseillers et enregistrés dans leur application de gestion des appels. Ces extractions bimensuelles ont été compilées pour former un panel qui a été anonymisé et dont l'unité d'observation est un conseiller téléphonique travaillant sur une campagne d'appel un jour donné. La performance est ainsi mesurée au niveau individuel quotidiennement. Ces données reflètent à la fois l'allocation du temps des conseillers dans les états du système telle qu'elle est décrite dans le schéma 1 et les issues de contact retranscrites par les conseillers au fur et à mesure de leur traitement des appels. La base de données couvre 21 semaines de référence avant l'expérimentation et les 16 semaines pendant lesquelles l'expérimentation a été mise en œuvre.

Sur la qualité de vie au travail des conseillers

La mesure de la qualité de vie au travail des conseillers téléphoniques a été réalisée à l'aide de questionnaires en ligne adaptés à chacune des campagnes téléphoniques et administrés à l'issue de chaque période d'observation (voir annexe n°1). Les conseillers devaient répondre aux questionnaires soit de leur domicile soit sur leur lieu de travail pendant leur temps de pause. Ils devaient répondre à autant de questionnaires que de campagnes téléphoniques sur lesquelles ils avaient travaillé la semaine précédant l'envoi des questionnaires, considérée comme la période de référence pour la mesure de la qualité de vie au travail. Chaque passation de questionnaire durait environ dix minutes.

Ces questionnaires recueillent des indicateurs concernant la description des campagnes, la satisfaction, le stress et la fatigue, l'utilisation des statistiques par les superviseurs et la perception de la surveillance. Chaque item est assorti d'une échelle de réponse de type Likert en 4 points (par exemple Je trouve la surveillance informatique excessive 1 « pas du tout d'accord » et 4 « tout à fait d'accord ») ou en 7 points (Dans l'ensemble sur la campagne d'appel XXX, estimez-vous que votre travail est stressant ? 1 correspondant à « pas du tout stressant » et 7 à « extrêmement stressant »).

Cinq passations ont été réalisées avant l'expérimentation à partir de la semaine 45 (en 2013) jusqu'à la semaine 4 (en 2014) et huit pendant l'expérimentation correspondant aux huit périodes d'observation du dispositif expérimental.

Collecte de données complémentaires

Nous retenons pour l'évaluation des résultats les campagnes qui ont été continuellement actives pendant les 37 semaines d'observation. Cette restriction permet d'analyser les performances et la qualité de la vie professionnelle liées à une partie homogène de l'activité du centre d'appels. Les données collectées avant le début de la période expérimentale sont utilisées afin d'avoir une image de la performance et de la qualité de vie au travail avant l'expérimentation et de procéder à des tests statistiques additionnels.

Nous étions également présents sur le terrain à chaque fin de séquence de quinze jours afin de conduire des entretiens individuels ou collectifs non directifs avec les conseillers téléphoniques et l'équipe de supervision. Ce travail qualitatif a alimenté un journal de bord qui contenait également les incidents survenus au cours de la période d'observation expérimentale. Cette méthodologie d'enquête complète l'analyse des résultats quantitatifs et constitue le support des interprétations de nos résultats.

3. Limites

La conjoncture de l'économie entre 2013 et 2014 a été difficile pour les entreprises sous-traitantes, en particulier celles qui travaillaient avec le secteur des télécommunications (Souquet, 2014). Les difficultés de l'entreprise pour renouveler ses contrats de sous-traitance ou pour trouver de nouveaux débouchés l'ont donc amenée à prendre des mesures de restriction budgétaire : les Tickets restaurant de ses salariés ont été suspendus pendant trois mois et une partie des locaux qu'elle occupait a été relouée à d'autres sociétés. Cette période a été mal vécue par les conseillers, parce qu'elle a augmenté un sentiment d'insécurité de l'emploi et elle a engendré une augmentation de l'absentéisme.

L'expérimentation, en elle-même, n'a pas été impactée par ces difficultés puisque son principe repose sur la comparaison d'une condition expérimentale et d'une condition contrôle. Les deux conditions sont impactées de la même manière par le contexte. Néanmoins, il y a eu d'autres répercutions, notamment sur le taux de réponses aux questionnaires administrés au cours de l'expérimentation. Si un salarié est absent sur une période donnée, nous ne pouvons pas recueillir sa performance ni son ressenti concernant sa qualité de vie au travail sur cette période. De plus lorsque les salariés sont tendus et inquiets quant à leur avenir, ils ont tendance à moins répondre aux questionnaires qualité de vie au travail. Cependant, il est rare qu'une entreprise ouvre ses portes à des chercheurs lorsqu'elle connaît une telle situation. Cela nous a permis d'implanter une expérimentation et d'en observer ses effets dans un contexte économique et social difficile.

V. Données

1. Les participants à l'expérimentation

Les participants à l'expérimentation sont les conseillers téléphoniques du centre d'appels. Avant leur entrée dans l'entreprise, la plupart d'entre eux étaient au chômage et ont bénéficié de la période de formation de quatre mois et demi financée par Pôle Emploi, avant leur contractualisation en emploi à durée indéterminée en 2012. Une seconde vague de recrutement a été effectuée en 2013, cette fois à durée déterminée pour des campagnes d'appels de court terme. Enfin, des apprentis ont été embauchés en alternance pendant quatre mois. Ces agents ont eu l'opportunité d'être embauchés en CDI.

Après avoir demandé leur consentement à participer à l'étude, 46 conseillers volontaires sur 58 potentiels ont choisi de participer, soit environ 80 % des employés du centre d'appel (voir tableau 1). La majorité des conseillers sont des femmes (69,9 %), et ils ont en moyenne 32,5 ans. Au début de l'expérimentation, 78,3 % des participants qui étaient en contrat à durée indéterminée travaillaient depuis 13,1 mois en moyenne, le reste des participants étaient en fin de formation et

ont signé leurs CDI au début de l'expérimentation. 28,3 % des conseillers travaillaient sur un seul type de campagne d'appels, et davantage sur des campagnes d'appels sortants.

Tableau 1 : Statistiques descriptives sur les conseillers téléphoniques du site expérimental

Taux ou moyenne (écart-type)	Participants		Non participants
	Total	Répondants au questionnaire QVT	
Part de femmes	69,6	85,7	60,0
Âge (en années)	32,5 (8,2)	35,4 (8,3)	33,6 (8,5)
Part de CDI	78,3	100	70,0
Part d'apprentis	21,7	0	15,0
Part de CDD	0	0	15,0
Ancienneté en CDI (en mois)	13,1 (4,45)	13,4 (3,0)	12,0 (4,78)
Part de conseillers travaillant sur un seul type de campagne téléphonique – part en appels entrants – part en appels sortants	28,3 4,4 23,9	7,1 0,0 7,1	40,0 20,0 20,0
Part de conseillers travaillant sur deux types de campagnes d'appels	71,7	92,9	60,0
Nombre de conseillers	46	14	20

Note : la deuxième colonne (QVT) correspond aux conseillers qui ont répondu aux questionnaires de qualité de vie au travail spécifiques aux campagnes d'appels.

Six superviseurs travaillaient sur la plateforme d'appels pendant la durée de l'expérimentation. Les superviseurs n'ont pas directement pris part à l'expérimentation, mais ils avaient ainsi en charge des conseillers de la condition expérimentale et des conseillers de la condition témoin toutes les deux semaines.

2. Les mesures de la performance

La performance dans un centre d'appels a de multiples facettes que l'on approche avec un jeu d'indicateurs plutôt qu'avec une mesure unique ce qui reflète la diversité des transactions de service (Gillet, Greenan et Le Gall, 2015). Nous retenons tout d'abord les indicateurs utilisés dans les contrats de service qui lient le centre d'appels sous-traitant et les donneurs d'ordres. Ces indicateurs sont définis au niveau d'une campagne dans les contrats de service, mais nous avons vu qu'ils étaient la plupart du temps utilisés pour évaluer la performance individuelle des conseillers. Les indicateurs du contrat de service contribuent ainsi à façonner les normes de travail sur le plateau téléphonique. Nous ajoutons à ces indicateurs d'autres mesures de la performance des conseillers qui sont propres à l'émission ou à la réception d'appels puisque le processus de production et les leviers de performances diffèrent fortement selon que l'appel est à

l'initiative ou non du conseiller. Nous avons donc identifié, pour chaque type de campagne d'appels, un ensemble de trois indicateurs.

a. Pour les appels entrants

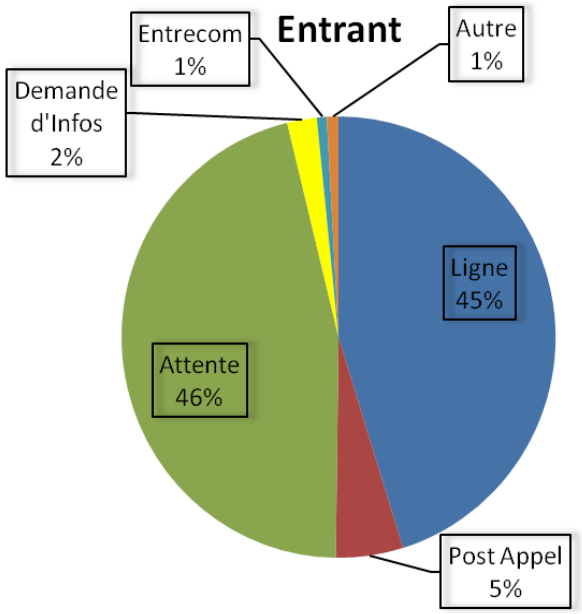
Le premier indicateur de performance en appel entrant est la durée moyenne de traitement (DMT). C'est l'indicateur le plus fréquemment utilisé dans les centres d'appels et négocié dans les contrats de sous-traitance. En principe, c'est le temps dont dispose chaque conseiller pour traiter le problème d'un appelant et pour remplir la fiche d'appels à l'issue de la résolution de la communication. L'objectif d'un contrat de service est d'optimiser l'externalisation des appels à la firme sous-traitante : le contrat fixe un objectif de durée d'appel qui permet de maintenir une file d'attente à un niveau raisonnable, sur lequel il existe également un seuil contractuel. En pratique, la taille de la file d'attente dépend de la taille de l'équipe de conseillers téléphoniques sur chaque campagne d'appel et sur le volume d'appels et sur la durée moyenne de traitement. Sur ces trois déterminants, seule la DMT est l'indicatrice sur laquelle les agents peuvent avoir un réel impact.

Le deuxième indicateur est la proportion de temps passé en post appel qui est également considéré comme un indicateur crucial pour la gestion de l'activité du plateau téléphonique. Le post appel est le temps qui est passé à traiter les derniers éléments d'un appel en éditant le fichier client pour notamment signaler la résolution d'un problème ou la mise à jour des informations de l'appelant (domicile, contact), ou divers éléments susceptibles d'aider les autres conseillers téléphoniques pour de futurs contacts. Cependant, les superviseurs ont tendance à vouloir réduire ce temps qui n'est pas dévolu à répondre aux appels : il est interprété par les employés du plateau téléphonique comme un temps prohibé source d'opportunisme pour couper le flux d'appel. C'est d'ailleurs à cet effet que la fenêtre « pop-up », « jaillissante », qui indique que « le temps de post appel est trop long » apparaît sur l'écran des conseillers à chaque fois que ce temps dépasse un seuil fixé.

Le dernier indicateur est la durée moyenne en attente des appelants. Puisque la satisfaction des clients dépend fortement de la durée de leur attente avant d'avoir un contact, cet indicateur est un élément important de la qualité de service. C'est pourquoi les contrats de sous-traitance définissent un niveau de satisfaction-service comme un seuil maximal d'attente avant qu'un appel soit traité par un conseiller. Comme ce temps en attente peut être mesuré pour chaque appel, il peut être également agrégé pour chaque campagne d'appels prise en charge par les conseillers téléphoniques. Cependant, il reflète davantage la performance sur la production globale que sur une performance individuelle.

Le travail des conseillers en réception d'appels est fortement contraint par la volumétrie journalière des appels entrants. Il est demandé aux conseillers de répondre vite aux appels et de traiter efficacement la demande d'un appelant, afin de ne pas engorger la file d'attente. Dans le même temps, le travail du superviseur est de gérer la file d'attente en allouant plus de conseillers à une campagne qui connaît un pic d'appels. Si la pression de la file d'attente est trop forte ou si le traitement d'un appel a été difficile, un moyen pour un conseiller de prendre un temps de repos est d'allonger son temps de post-appel. Le graphique 1 donne la décomposition du temps de production dans les différents états du système en réception d'appels. Nous constatons que les temps différents de ceux en attente et en ligne sont très comprimés en réception d'appels. La majeure partie du temps est passé en ligne et en attente avec une part de temps d'ampleurs comparables dans les deux états (respectivement 45 % et 46 % en moyenne).

Graphique 1 : Décomposition du temps de production en réception d'appels



b. Pour les appels sortants

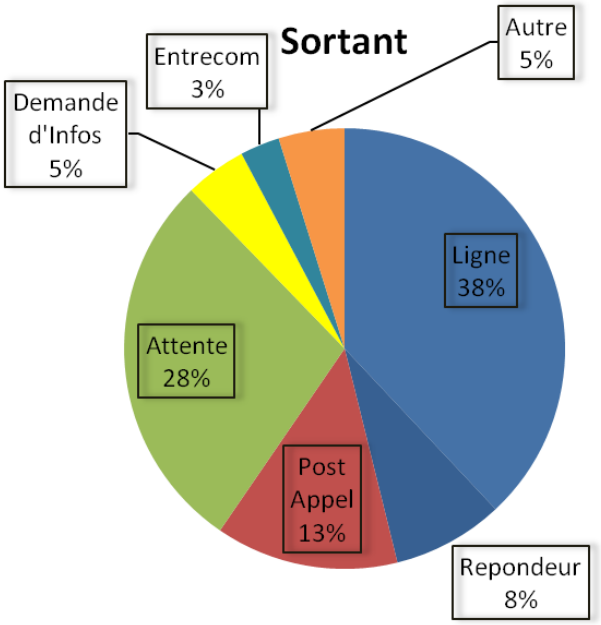
Le premier indicateur est le taux de concrétisation (ou taux de conversion) qui mesure la proportion d'accords conclus pendant un jour donné sur le nombre de contacts argumentés, c'est-à-dire lorsque la communication a abouti à une discussion qui a permis de dérouler le script de vente d'un produit, quelle que soit la réponse du prospect. Il mesure donc l'effort fourni par un conseiller pour obtenir un accord. Il est mesuré au niveau de chaque campagne d'appel, mais ce n'est pas forcément un indicateur contractualisé.

Le deuxième indicateur est la part du temps passé en post appel. Il est choisi pour les mêmes raisons que celles exposées pour les campagnes d'appels entrants. Cependant le processus de production étant moins contraint pour les campagnes d'appels sortants qu'en appels entrants, étant donné que le flux d'appel n'impose pas le rythme du travail, cet indicateur est davantage pris en compte pour les campagnes d'appels sortants parce qu'il y a plus de risque de voir les conseillers prendre plus de temps entre chaque appel.

Enfin le troisième indicateur est la cadence d'appels, calculée à partir du nombre de contacts à l'heure. Cet indicateur mesure l'intensité du travail des conseillers téléphoniques. Par ailleurs, l'utilisation du mode prédictif influence cet indicateur parce qu'il oblige les conseillers à contacter plus de prospects à la chaîne.

Le travail des conseillers en émission d'appels est contraint essentiellement par la cadence d'appels à fournir, dans le but de décrocher des accords sur une liste prédéfinie de prospects (le fichier client). Le conseiller appelle les individus en composant manuellement leurs numéros ou bien le système informatique effectue une numérotation automatique (appels prédictifs). Le graphique 2 donne la même décomposition que le graphique 1 pour les campagnes d'appels sortants. Les conseillers passent en moyenne 38 % de leur temps en ligne et 28 % de leur temps en attente. Comme l'appel est à l'initiative du conseiller, on peut être surpris par l'existence d'un temps en attente. Celui-ci est généré par les problèmes des fichiers d'adresses, la disponibilité des appelés et le processus d'appel automatique pour le mode prédictif. Le travail en émission d'appels est plus répétitif, car il démarre toujours par la même séquence destinée à contacter l'appelé. Elle débouche la plupart du temps sur un téléphone qui sonne dans le vide ou sur le message d'un répondeur. On observe également que les temps autres que les états en ligne et en attente sont plus importants en appels sortants qu'en appels entrants. Cela signale en creux la pression de la file d'attente dans le travail en émission d'appels et le rôle central de l'engagement du salarié dans la fixation du rythme de l'émission d'appels.

Graphique 2 : Décomposition du temps de production en émission d'appels



3. Les mesures de la qualité de vie au travail

Plusieurs indicateurs étaient mesurés à l'aide des questionnaires de qualité de vie au travail, mais les plus importants concernaient la satisfaction au travail, le niveau de stress et la perception de l'intensité de la surveillance.

La satisfaction au travail est mesurée à l'aide d'une échelle de Lickert à quatre points. Nous avons construit cet indicateur à partir d'un faisceau de dix items pour les campagnes d'appels entrants et sortants : le niveau de service fourni ; les interactions avec les clients ; la qualité du script ; le soutien reçu de la part des collègues ; le soutien reçu de la part des superviseurs ; la reconnaissance du travail accompli par les superviseurs ; les sessions de formation pour les campagnes d'appels ; la programmation des emplois du temps ; les primes ; le travail sur les campagnes en général. Un item additionnel à propos de la qualité du fichier prospect est ajouté sur la mesure de la satisfaction sur les campagnes d'appels sortants[16].

Le stress au travail est mesuré sur une échelle de Lickert à sept points : « en général, pensez-vous que votre travail sur la campagne d'appel X est stressante ? » « 1 : pas stressante du tout ; 7 : extrêmement stressante ». Nous avons également mesuré si le stress est incontrôlable sur une échelle de Lickert à sept points.

Enfin, comme les superviseurs sont affectés à une campagne d'appels donnés, la perception de la surveillance est susceptible de varier d'un conseiller téléphonique à un autre et d'une campagne à l'autre. C'est pourquoi nous avons demandé aux conseillers de mesurer leur ressenti à propos de l'intensité de la surveillance pour chaque campagne d'appels en utilisant deux questions : la

[16] L'alpha de Cronbach mesurant l'adéquation de l'agrégation des indicateurs pour l'indicateur de satisfaction est supérieur à 0,80.

première s'attache à questionner l'excessivité de la surveillance sur une échelle de Lickert à quatre points (« Je trouve que la surveillance électronique est excessive. 1 : totalement en désaccord ; 4 : totalement d'accord) ; la seconde questionne la permanence de la surveillance sur une échelle de Lickert à quatre points.

Nous avons subi une attrition dans les réponses des conseillers téléphoniques. En effet, un ressenti de lassitude des agents a été observé sur le terrain, notamment à cause de la répétitivité des questions posées et de la succession des questionnaires. Un moyen utilisé pour augmenter le taux de réponse a été de raccourcir la longueur des questionnaires pour se concentrer uniquement sur les items les plus pertinents pour notre étude.

4. Stratégie d'identification

Pour notre stratégie d'identification des effets de la condition expérimentale, c'est-à-dire de la réduction de l'intensité de la surveillance électronique sur la performance des conseillers téléphoniques et sur leur qualité de vie au travail, nous nous concentrons sur douze campagnes d'appels téléphoniques qui ont été actives durant les trente-sept semaines d'observations et pour les quarante-six participants. Ce cadre d'analyse permet de former 234 combinaisons de conseillers-campagnes d'appels.

Cependant, les bases de données mesurant les indicateurs de performance et de qualité de vie n'ont pas les mêmes structures. Nous avons donc utilisé différentes méthodes statistiques appliquées aux caractéristiques des données, mais qui permettent de formuler des résultats comparables.

a. Sur la performance

Nous avons collecté un ensemble exhaustif d'informations journalières sur les indicateurs de performance des conseillers téléphoniques. La base de données contient ainsi plus de 9 000 enregistrements quotidiens individuels sur l'ensemble des campagnes d'appels.

Pour évaluer l'effet moyen de la condition expérimentale, l'assignation aléatoire des individus entre le groupe traité et le groupe témoin tous les quinze jours, nous assure la constitution d'un groupe traité et d'un groupe témoin. Par conséquent, nous pouvons mesurer sans biais l'estimateur de l'effet causal suivant (Rubin, 1974), noté γ :

$$\gamma = E(Y_T) - E(Y_C)$$

où $E(Y_T)$ $(E(Y_C))$ est le résultat moyen du groupe traité (respectivement non traité). Nous pouvons estimer cet indicateur par un modèle de régression linéaire qui tient également compte de la dimension temporelle et de la dimension du travail spécifique à chaque campagne d'appels.

Le modèle que nous utilisons est le suivant :

$$Y_{i_{j_c}} = \alpha_1 . T_j + \gamma . G_{i_j} + \alpha_{i_c} + \beta + \epsilon_{i_{j_c}}$$

où $Y_{i_{j_c}}$, qui mesure l'indicateur de performance quotidien pour un conseiller téléphonique, est expliqué par une variable dichotomique temporelle T_j qui indique le début de la période expérimentale et par la variable dichotomique G_{i_j} qui indique l'appartenance au groupe de test. Parce que nous utilisons un panel de conseillers travaillant sur des campagnes d'appels données sur des jours donnés, des effets fixes α_{ic} peuvent être inclus dans la régression pour contrôler l'hétérogénéité inobservable invariante dans le temps au niveau des conseillers travaillant sur les campagnes d'appels.

Ce modèle de régression mesure l'effet causal du traitement, en contrôlant les différences invariantes dans le temps dans les mesures de performance pour chaque conseiller téléphonique traitant une campagne donnée. Nous testons l'hypothèse nulle, selon laquelle les différences de moyennes entre les deux conditions de l'expérimentation ne sont pas significativement différentes de 0, avec un test de Student.

b. Sur la qualité de vie au travail

Notre base de données sur la qualité de vie au travail est structurée différemment, car nous ne collectons pas d'informations quotidiennes comme pour les indicateurs de performance. De plus, nous avons fait face à un problème d'attrition pendant l'expérience. Nous avons également mentionné le fait que la situation économique mondiale des centres d'appels devenait de plus en plus difficile durant l'expérimentation. En conséquence, 429 questionnaires ont été entièrement remplis, soit 36 % des observations attendues, correspondant aux réponses de quatorze conseillers traitant différentes campagnes d'appel (voir le tableau 1, colonne QVT). Nous appliquons donc une autre méthode statistique, adaptée aux caractéristiques des données, qui donne des résultats comparables à ceux obtenus par l'analyse économétrique.

Nous avons procédé à une mise en correspondance des réponses des agents d'appel entre les moments où ils ont été affectés au groupe test et les moments où ils ont été affectés au groupe témoin pour la même campagne. Par exemple, si un agent était dans le groupe test la première quinzaine et dans le groupe témoin dans les autres, nous avons couplé ses réponses du premier questionnaire à ses sept autres réponses. En conséquence, nous obtenons un échantillon apparié sur 360 questionnaires que nous avons utilisé pour calculer les différences de niveau individuelles entre les deux conditions et les différences de groupe moyennes correspondantes. Nous utilisons un test de Student pour échantillon apparié pour vérifier si les différences observées sont significatives.

VI. Résultats

1. Sur les campagnes d'appels entrants

a. Les indicateurs de performance individuelle

La période précédant l'expérimentation pour les campagnes d'appels entrants contient 2 799 observations journalières conseiller-campagne. Les huit assignations aléatoires ont permis de constituer un groupe de traitement (1 183 observations) et un groupe témoin (1 206 observations). Les trois premières colonnes du tableau 2 donnent les statistiques descriptives des indicateurs de performance dans la période préexpérimentale et pendant l'expérimentation.

Tableau 2 : Effets de la condition expérimentale sur les indicateurs de performances pour les campagnes d'appels entrants

Indicateurs	Avant l'expérimentation	Expérimentation		Avec un effet fixe conseiller-campagne	
		Groupe test	Groupe témoin	Différence	t
Durée moyenne de traitement (en secondes)	221,15 (143,60)	211,62 (120,49)	224,68 (143,87)	2,23 (4,99)	0,446
Part du temps passé en post-appel (sur le temps de production)	6,54 (7,88)	7,34 (9,07)	6,81 (9,29)	0,52 (0,35)	1 510
Durée moyenne en attente (en secondes)	45,19 (59,81)	45,60 (60,84)	45,97 (63,57)	1,54 (2,18)	0,710
Nombre d'observations	2799	1183	1206	5188	

Note : les écarts-types sont donnés entre parenthèses dans les trois premières colonnes. L'erreur-type est présentée entre parenthèses dans la quatrième colonne. Les différences ne sont pas significatives au seuil de 10 %.

La colonne intitulée « différence » donne l'impact causal de la condition expérimentale estimé à partir de notre modèle de régression. Les résultats montrent que notre condition expérimentale n'a pas eu d'effet significatif sur les indicateurs de performance pour les campagnes d'appels entrants. En tenant compte de l'hétérogénéité inobservable entre les conseillers téléphoniques travaillant sur une campagne d'appel donné, il n'y a pas de différence significative entre le groupe traité et le groupe témoin.

b. Les indicateurs de qualité de vie au travail

Le tableau 3 montre les résultats à propos de la qualité de vie au travail des conseillers téléphoniques lorsqu'ils travaillaient sur les campagnes d'appels entrants.

Tableau 3 : Indicateurs de qualité de vie au travail sur les campagnes d'appels entrants

Indicateurs	Avant l'expérimentation	Expérimentation		Différence	t
		Groupe test	Groupe témoin		
Satisfaction	3,03 (0,53)	2,66 (0,58)	2,78 (0,60)	**-0,12**** **(0,37)**	-3 734
Stress	3,58 (1,55)	4,82 (1,15)	4,16 (1,34)	**0,66**** **(1,53)**	4 912
Le facteur de stress est incontrôlable	3,79 (2,09)	5,02 (1,21)	4,94 (1,45)	0,08 (1,31)	0,739
Surveillance excessive	2,62 (1,09)	3,28 (0,97)	3,11 (0,87)	**0,17*** **(0,86)**	2 164
Surveillance non permanente	2,58 (1,13)	2,13 (1,16)	2,32 (0,94)	-0,19 (1,10)	-1 919
Nombre d'observations	77	129	129	258	

Note : les écarts-types sont donnés entre parenthèses dans les trois premières colonnes. L'erreur-type est présentée entre parenthèses dans la quatrième colonne. Les différences sont significatives : *p < 0.05, **p < 0.01. Six observations sont manquantes sur les items de surveillance.

Lorsqu'ils étaient dans la condition expérimentale, les conseillers ont ressenti que le travail était moins satisfaisant en moyenne et plus stressant. De plus, les conseillers ont exprimé un sentiment d'être surveillé de façon plus permanente et ils ont ressenti une surveillance plus excessive par rapport à la situation où la surveillance électronique était maintenue à un niveau inchangé.

2. Les campagnes d'appels sortants

a. Les indicateurs de performance individuelle

Les statistiques descriptives avant et pendant l'expérimentation sont présentées dans les trois premières colonnes du tableau 4. La période préexpérimentale contient 1615 observations journalières. Après huit assignations aléatoires, le groupe traité contient 1082 observations tandis que le groupe témoin contient 1264 observations. L'activité des campagnes d'appels sortants a davantage augmenté par rapport aux campagnes d'appels entrants pendant la période expérimentale : le nombre d'observations en appel entrant était le double de celles des campagnes d'appels sortants, alors qu'elles sont similaires pendant l'expérimentation. Cela s'explique par le fait que la plateforme téléphonique a réduit le nombre de conseillers sur chaque campagne d'appels entrants.

Tableau 4 : Les effets de la condition expérimentale sur les indicateurs de performance en appels sortants

Indicateurs	Avant l'expérimentation	Expérimentation		Avec un effet fixe individu-campagne	
		Groupe test	Groupe témoin	Différence	t
Taux de concrétisation (part des accords sur les contacts argumentés)	41,67 (38,16)	34,41 (40,52)	32,94 (39,28)	**2,49* (1,42)**	1 760
Part du temps passé en post-appel (sur le temps de production)	13,91 (13,55)	11,86 (15,95)	10,82 (14,44)	**-0,92** (0,392)**	-2 350
Cadence (nombre d'appels à l'heure)	24,11 (13,42)	15,07 (11,27)	15,98 (11,27)	-0,53 (0,37)	-1 510
Nombre d'observations	1615	1082	1264	3961	

Note : les écarts-types sont donnés entre parenthèses dans les trois premières colonnes. L'erreur-type est présentée entre parenthèses dans la quatrième colonne. Les différences sont significatives : *p < 0.05, **p < 0.01.

Les résultats du modèle de régression qui estime l'effet causal de la condition expérimentale sont présentés dans la quatrième colonne. Ils montrent que le groupe traité a eu un gain de taux de concrétisation de 2,5 points de pourcentage comparé au groupe témoin. Le groupe traité a également diminué le taux de temps passé en post-appel de 0,92 point de pourcentage. Ces résultats témoignent d'un gain d'efficacité des conseillers téléphoniques lorsqu'ils étaient dans un environnement de travail où la surveillance électronique de leur effort était réduite.

b. Les indicateurs de qualité de vie au travail

Le tableau 5 présente les résultats de l'expérimentation pour les campagnes d'appels sortants.

Tableau 5 : Indicateurs de qualité de vie au travail sur les campagnes d'appels sortants

Indicateurs	Avant l'expérimentation	Expérimentation		Différence	t
		Groupe test	Groupe témoin		
Satisfaction	2,74 (0,63)	2,56 (0,54)	2,62 (0,48)	-0,06 (0,41)	-1 100
Stress	3,80 (1,93)	3,92 (1,69)	3,51 (1,56)	**0,41* (1,40)**	2 043
Le facteur de stress est incontrôlable	4,26 (2,02)	4,80 (1,96)	4,08 (1,82)	**0,72** (1,84)**	2 809
Surveillance excessive	2,53 (1,08)	2,30 (1,09)	2,02 (1,19)	**0,28** (0,70)**	2 824
Surveillance non permanente	2,50 (1,06)	2,66 (1,15)	2,76 (1,04)	-0,10 (1,05)	-0,670
Nombre d'observations	133	51	51	102	

Note : les écarts-types sont donnés entre parenthèses dans les trois premières colonnes. L'erreur-type est présentée entre parenthèses dans la quatrième colonne. Les différences sont significatives : *p < 0.05, **p < 0.01. Avant l'expérimentation, six observations sont manquantes sur l'item facteur de stress est incontrôlable ; deux sont manquantes sur les items de surveillance. Pendant l'expérimentation, une observation est manquante sur les items de surveillance, deux sur l'item de stress.

Il ne semble pas y avoir de différences significatives en moyenne au sujet de la satisfaction au travail. Cependant, lorsqu'ils étaient dans la condition expérimentale, les conseillers téléphoniques ont ressenti davantage de stress et ce facteur de stress était davantage incontrôlable. La surveillance était perçue comme plus excessive pour les conseillers sous la condition expérimentale.

3. Synthèse

Pour résumer, les résultats de l'expérimentation supportent l'Hypothèse 2. La manipulation de l'intensité de la surveillance dans la condition expérimentale n'a pas eu de conséquences négatives sur la production : en ce qui concerne les campagnes d'appels entrants, il n'y a pas eu de différences significatives sur les indicateurs de performance entre le groupe traité et le groupe témoin ; en ce qui concerne les campagnes d'appels sortants, nous avons détecté un taux de concrétisation plus important et un temps passé en post appel inférieur. Par conséquent, la condition expérimentale n'a pas détérioré la performance, et cela n'a pas conduit les conseillers à adopter des comportements opportunisme de relâchement de leur effort lorsque la surveillance était réduite. Cela conduirait à valider l'Hypothèse 1 selon laquelle le niveau de surveillance était certainement excessif.

En revanche, les résultats concernant la qualité de vie au travail sont plus surprenants et ils ne supportent pas l'Hypothèse 3 : la condition expérimentale n'a pas permis d'améliorer la qualité de

vie au travail et au contraire cela l'a même détériorée. De plus, de manière paradoxale, le dispositif expérimental qui consistait à réduire le niveau de surveillance électronique a eu pour conséquence un ressenti plus excessif de la surveillance. Il existerait donc d'autres déterminants de la qualité de vie au travail qui auraient modéré l'impact de la condition expérimentale.

VII. Interprétations

Si nos conclusions à propos de la qualité de vie au travail ne sont pas en adéquation avec les principaux résultats obtenus dans la littérature, nous avons néanmoins noté dans notre revue des études l'ambiguïté structurelle de la surveillance et nous avons listé les facteurs modérateurs du lien entre surveillance et qualité de vie au travail : la difficulté de la tâche et le paramétrage du système de surveillance. Cependant, l'expérimentation a eu lieu dans un contexte de travail réel et non dans un laboratoire. Par conséquent, les caractéristiques du système de surveillance nous étaient données, et malgré nos sessions d'observations sur le terrain et nos entretiens avec les employés, une partie de son fonctionnement est resté dissimulé. Nous n'avons ainsi pas eu la totale maîtrise du système de surveillance. Néanmoins, grâce à notre procédure d'assignation aléatoire des conseillers téléphoniques, nous sommes plutôt sûrs que l'effet causal que nous avons mesuré est indépendant de tout autre facteur objectif que celui que nous avons manipulé. Par exemple, ces différences ne peuvent pas être attribuées à une variation de la complexité de la tâche où à une conception des primes parce que les deux groupes sont strictement équivalents en ce qui concerne ces critères.

Le fait que notre condition expérimentale accroît le sentiment de surveillance est un paradoxe. Comme nous ne pouvons pas tester directement l'Hypothèse 1, nous sommes entrés plus loin dans la « boîte noire » de l'intensité de la surveillance. Nous avons donc utilisé les entretiens de terrain que nous avons conduits avec les employés pour comprendre comment ils avaient réagi aux changements des paramètres de leurs logiciels. Deux explications potentielles émergent de ces analyses liées aux représentations des acteurs et à leur utilisation du système d'évaluation de leur performance : la première est reliée à l'idée selon laquelle les indicateurs de performance en temps réel contribuent à structurer le sentiment d'auto-efficacité et de contrôle des conseillers téléphoniques ; la seconde est reliée aux représentations des superviseurs à propos de leur travail.

1. Une déstabilisation des conseillers téléphoniques

Nos résultats nous amènent dans un premier temps à nous interroger sur le rôle joué par les outils de mesure de la performance. Pour l'expliquer, nous nous référons à la « théorie de la fixation des objectifs » de Locke et Latham (1990), dans laquelle un objectif fixé maintiendrait voire augmenterait la performance. Les indicateurs de performance des plateformes

téléphoniques sont ainsi conçus comme des objectifs fixés par l'organisation auxquels chaque employé doit parvenir en fournissant un effort. L'efficacité de l'atteinte de ces objectifs dépend de plusieurs facteurs : l'acceptation de ces objectifs, les difficultés et les spécificités pour les atteindre, l'émission du *feedback* et la mise en place par les collaborateurs eux-mêmes de leurs propres objectifs.

Dans la condition expérimentale, les conseillers ne pouvaient plus accéder à certains retours sur leurs performances. Ils ont eu accès à des statistiques simples (nombre d'appels, d'accord…), mais pas à des statistiques complexes (y compris celles requises par le client comme la durée moyenne de traitement). Les conseillers connaissaient les objectifs, mais n'ont plus eu accès à des informations spécifiques sur leurs performances individuelles. Or, la comparaison entre l'objectif fixé et sa réalisation permet aux agents de réguler leur comportement (Kluger et DeNisi, 1996) et il semble que les effets détériorant de la qualité de vie au travail d'un système de surveillance électronique peuvent être atténués lorsque ce dernier est associé à des critères de performance clairs et à un *feedback* régulier (Bates et Holton, 1995 ; Holman et al., 2002). Dans une expérience de laboratoire menée auprès de 165 étudiants de premier cycle, Alder et Ambrose (2005) montrent l'intérêt du *feedback* dans un système de surveillance électronique : ils ont constaté qu'un *feedback* spécifique des superviseurs, ressenti comme constructif, affecte directement l'équité perçue du suivi électronique des performances et affecte indirectement la performance et la satisfaction en termes de quantité et de qualité. Inversement, Earley (1988) a montré que le *feedback* fourni par un ordinateur permettait d'obtenir de meilleures performances qu'un *feedback* fourni par un superviseur. Kluger et DeNisi (1996) soulignent que les effets du *feedback* sur la performance sont variables et que lorsque les personnes reçoivent un commentaire négatif, elles ont tendance à accroître leurs efforts.

Dans notre étude, la suppression de l'accès aux statistiques a donc augmenté le niveau d'incertitude des conseillers. Ils n'ont plus eu d'accès à l'information pour diriger leur comportement. Dans l'incertitude, il semble que les conseillers ont augmenté leur niveau d'effort, ce qui a conduit à une augmentation de la performance.

Par conséquent, les conseillers téléphoniques auraient été déstabilisés par la suppression de l'accès à leurs indicateurs de performance en temps réel et par la suppression de la file d'attente. La perception de ce changement aurait été vécue de façon hétérogène parmi les conseillers lorsqu'ils travaillaient sous la condition expérimentale : dans les entretiens, une partie d'entre eux ont pointé le fait que la possibilité de vérifier ces indicateurs contribuait à un sentiment de contrôle et d'auto-efficacité. Dans un contexte où le produit issu du service rendu est immatériel, la variété des indicateurs de performance permet de matérialiser le travail réalisé et de maintenir la motivation lorsqu'ils s'ennuient, mais également de leur procurer un sentiment d'assurance alors qu'ils pensent être en dessous des seuils d'objectifs fixés, et de réguler leur effort pendant les pics

d'activité. Pour ces conseillers, les indicateurs de performance et la file d'attente sont des ressources et leur travail s'appuie sur eux. Le fait d'avoir désactivé l'accès à une partie de l'information dans une situation financière et sociale tendue a vraisemblablement accru le sentiment d'une insatisfaction dans le travail.

Cela a pu produire une incitation à produire un niveau d'effort supérieur dans la condition expérimentale afin d'être certain de maintenir leur niveau de performance (Bartels et Nordstrom, 2012) tandis que dans le même temps, leur satisfaction diminuait et leur stress augmentait. Les résultats de performance confirment cette intuition dans les campagnes d'appels sortants. En effet, l'engagement des conseillers dans le processus de production est un élément central de la performance pour ce type d'activité. En revanche, cette interprétation ne permet pas d'expliquer le sentiment d'un accroissement de l'intensité de la surveillance.

2. Un surajustement des superviseurs

Pour expliquer le résultat paradoxal à propos de la perception d'une surveillance plus excessive sur les deux types de campagne d'appels et plus permanente pour les campagnes d'appels entrants, alors que la condition expérimentale devait réduire l'intensité de la surveillance, nous nous sommes tournés du côté du comportement des superviseurs. Malgré la suppression de certains indicateurs de performance, les superviseurs auraient ainsi développé d'autres moyens de surveillance qui s'apparenterait à un contrôle physique plus accru du travail, qu'à une véritable évaluation des objectifs fixés, ce qui contribue à réduire le niveau de confiance dans le travail des conseillers.

Selon Chouanière et al. (2011), un comportement plus contrôlant des superviseurs est un facteur qui a des effets négatifs sur le bien-être des employés des centres d'appels.

Griffith (1993), bien qu'ayant montré qu'il n'y avait pas de différence de performance entre la surveillance électronique et celle d'un superviseur, observe des pointes de performance en cas de présence physique du superviseur alors que les performances sous surveillance informatique restent stables. Pour Griffith (1993), si la performance est plus constante dans la condition de surveillance électronique peut être due à l'incertitude concernant l'évaluation de la performance : lorsque le superviseur n'est pas présent, le conseiller téléphonique sait qu'il n'est pas surveillé alors il maintient une performance constante.

Stanton et Sarkar-Barney (2003) ont montré que la performance qualitative variait selon le type de surveillance contrairement à la performance quantitative. La performance qualitative est meilleure dans un groupe sous surveillance informatique et dans un groupe non surveillé par rapport au groupe de surveillance directe humaine, ce qui suggère un plus grand effet de facilitation dans ce dernier groupe.

En outre, selon Stanton (2003), il y aurait une sorte d'accoutumance à la surveillance électronique, c'est-à-dire que d'une part les conseillers téléphoniques ont intégré dans leur façon de travailler l'intensité de la surveillance pratiquée par les superviseurs et y sont habitués et d'autre part c'est une partie essentielle de l'activité des superviseurs. De sorte que les superviseurs, en utilisant des moyens de surveillance détournés, doivent l'avoir rendu plus visible. Cela expliquerait l'augmentation de certains indicateurs de performance et la diminution de la qualité de la vie professionnelle dans des conditions expérimentales.
Enfin, si nous nous référons aux théories sur de l'effet d'éviction de la motivation intrinsèque (Frey et Jengen, 2001), nous pouvons suggérer que les pratiques de surveillance alternatives utilisées dans la condition expérimentale ont envoyé un signal de méfiance beaucoup plus important que les pratiques habituelles. En effet, lorsque les superviseurs doutaient du comportement d'un conseiller, ce dernier se sentait immédiatement pris pour cible et la méfiance était déterminée à devenir rapidement un constat partagé sur le plateau téléphonique. Les superviseurs ne pouvaient plus étayer leur soupçon en se référant au temps passé dans un état du système qu'ils considèrent comme improductif. Cette explication du paradoxe de la surveillance pourrait également mettre en évidence un stress et une insatisfaction accrus à l'égard des campagnes d'appels.

Nous nous attendions à ce que les superviseurs réallouent leur temps passé à vérifier les indicateurs de performance en temps réels vers des périodes de gestion des incidents techniques, de coordination du plateau téléphonique ou de tâches permettant d'améliorer la confiance en procurant du soutien, du *feedback* et de la formation. Les entretiens ont permis de révéler que ces stratégies d'adaptation au changement organisationnel ne se sont pas vérifiées. En réalité, les superviseurs perçoivent la surveillance comme une partie intégrante de leurs activités et par conséquent ils ont réagi à la perte de leurs outils de surveillance individuelle en temps réel, en particulier le chronométrage des différents états des conseillers, en utilisant d'autres moyens pour maintenir leur niveau de surveillance : ils ont davantage utilisé la prise de contrôle à distance des postes de travail des conseillers pour vérifier leurs durées d'appels ; ils ont été davantage à l'écoute des conversations pour vérifier la justesse des suivis de scripts ; et ils ont eu une présence physique sur le plateau téléphonique plus fréquente ou moins prévisible. Les superviseurs des centres d'appels ont ressenti une forte pression de la part de la hiérarchie et des clients pour atteindre les indicateurs de performance requis. Ils ont tendance à avoir plus de contrôle sur les conseillers qui ne paraissent pas motivés et il semble que pour certains, la meilleure façon d'obtenir la performance attendue est de contrôler leur travail. Ces pratiques de contrôle, plus que d'évaluation des objectifs atteints, ont été perçues comme plus intrusives de la part des conseillers que celles qui étaient alors utilisées avant l'expérimentation. Par conséquent, même si les

superviseurs avaient objectivement moins d'informations en temps réel sur les performances quand les conseillers téléphoniques étaient dans le groupe test que lorsqu'ils étaient dans le groupe témoin, parce que les pratiques alternatives de surveillance se révélaient beaucoup moins discrètes, les agents ont perçu la surveillance comme excessive et permanente.

Dans notre expérience, on peut supposer que l'augmentation des performances et la détérioration de la qualité de vie au travail sont dues à la présence physique d'un superviseur sur la plateforme d'appels. Cependant, dans notre condition expérimentale, même si les superviseurs n'étaient pas présents sur le plateau d'appels, ils ont toujours eu accès à des états dans lesquels étaient les conseillers sans la durée. Au cours des entretiens, les superviseurs nous ont expliqué que lorsqu'ils ont accès à la durée, ils attendent habituellement 10 minutes avant d'agir. Les conseillers téléphoniques savent que lorsqu'ils sont connectés à un état non productif ou à un état productif pendant plus de 10 minutes, un superviseur viendra à leur rencontre pour les réprimander ou les aider. Parfois, les conseillers se connectent à des états improductifs afin de prendre des pauses. Il existe des ruptures cachées qui pourraient renvoyer aux stratégies de résistance décrites par Bain et Taylor (2000). Sans informations sur la durée, les superviseurs peuvent intervenir à tout moment lorsqu'un agent d'appel est dans un état non productif. Leurs comportements deviennent imprévisibles pour les agents d'appel. Cela pourrait exacerber leur sentiment d'être surveillé et leur appréhension de l'évaluation.

Les intensités de surveillance basées sur la quantité d'informations disponibles dans le système et sur la base du niveau de pression exercée par les superviseurs à travers leurs pratiques de surveillance doivent ainsi être distinguées. La réaction des superviseurs à la condition expérimentale indique une cause organisationnelle de surveillance excessive. Le fait que les superviseurs se sentent personnellement responsables de la conformité aux normes établies dans les contrats de niveau de service est susceptible d'expliquer pourquoi ils ont utilisé des stratégies d'adaptation pour maintenir l'intensité de surveillance des agents d'appel dans le groupe test. Ils ont essayé de maintenir l'intensité du suivi parce qu'ils se sentaient personnellement responsables des mesures de performance des campagnes d'appels qui sont rapportées aux entreprises clientes de la plateforme téléphonique. La pression des donneurs d'ordre les conduirait à devenir plus contrôlants et moins orientés vers le soutien à l'autonomie. Cette pression est d'autant plus forte que la conformité à la norme fixée dans le contrat de service est déterminante pour sauvegarder l'emploi et sécuriser la survie du centre d'appels. Il y aurait donc eu un transfert de l'intensité de la surveillance qui a été perçue comme plus directe et plus visible.

Enfin, lors des derniers entretiens collectifs sur le terrain, les superviseurs ont décrit davantage leur rôle dans la gestion du plateau et plus particulièrement dans le rapport à leurs équipes. Les principes de respect de l'équité et de rapport à la hiérarchie sont des moteurs importants de la confiance accordée par les superviseurs. Ainsi, à partir du moment où les superviseurs ont lâché

la bride sur la surveillance en temps réel, ils ont estimé que l'évaluation quotidienne ou hebdomadaire était plus importante. Ils ont ainsi été plus exigeants dans l'effort agrégé que dans l'effort permanent et cette contrainte a été mal perçue par les conseillers qui ont eu des injonctions contradictoires : plutôt que de fournir un effort permanent, ils ont dû fournir un effort plus moyen. Et les superviseurs n'ont pas bien véhiculé ce message aux individus. Ce constat permet en partie de valider l'hypothèse sur la confiance et l'équité dans le rapport d'évaluation.

VIII. Conclusion

1. Une révélation détournée des effets néfastes de la surveillance électronique

Notre expérimentation de terrain a été conçue pour analyser les conséquences d'une réduction de l'intensité de la surveillance électronique sur la performance de conseillers téléphoniques et sur leur qualité de vie au travail. Nous avions supposé que l'intensité de la surveillance était susceptible d'être excessive dans un centre d'appels externalisé pour deux raisons principales. Premièrement, la numérisation des indicateurs de performance les avait rendus invisibles et deuxièmement, le transfert des contraintes de contrat de sous-traitance de service sur la relation d'emploi était intégré dans l'application du pilotage par les superviseurs.

Nous avons constaté que l'impact causal de notre condition expérimentale était une augmentation des performances dans les campagnes d'appels sortants, une performance stable dans les campagnes d'appels entrants et une qualité de vie au travail dégradée pour les deux types de campagnes. Si l'augmentation du stress et la diminution de la satisfaction au travail étaient inattendues, nous avons également trouvé un paradoxe de la surveillance : notre condition expérimentale, qui visait à diminuer l'intensité de la surveillance électronique des performances, entrainait un sentiment plus important de surveillance excessive. Nous avons donc revisité nos hypothèses en analysant les entretiens que nous avons menés sur le terrain avec des conseillers téléphoniques et des superviseurs, avant et pendant la période expérimentale. L'examen de ce matériau d'analyse sur le terrain à la lumière de nos résultats expérimentaux a révélé le fonctionnement réel du système de surveillance électronique de la performance.

Dans un système de surveillance électronique, la supervision se fait à travers le filtre des applications informatiques des superviseurs et des conseillers téléphoniques qui traitent les informations en temps réel sur les performances. Ces nouveaux outils peuvent changer le moment de la surveillance. Au lieu de se produire à intervalles de temps discrets, la surveillance

peut se produire en temps continu. Dans notre condition expérimentale, nous avons modifié les paramètres de l'application des superviseurs et des agents d'appel en supprimant les informations individuelles en temps réel sur les performances. Nous avions l'idée que cela réduirait la fréquence de la surveillance et libérerait ainsi du temps que les superviseurs utiliseraient pour des activités de coordination ou d'amélioration de la confiance comme le *feedback* ou la formation. Ce n'est pas ce qui est arrivé. Les superviseurs ont réagi à la perte d'informations sur les performances individuelles des conseillers dans le groupe traité en utilisant d'autres pratiques de surveillance plus explicites afin de maintenir le niveau d'intensité de surveillance. Les conseillers téléphoniques dans le cadre de la condition expérimentale ont estimé que la surveillance était plus excessive ou plus permanente parce que ces pratiques étaient plus pressantes que la surveillance assistée par ordinateur.

La littérature sur les coûts cachés de la surveillance montre que les employés réagissent à la surveillance et que parfois ces réactions contrecarrent les objectifs initiaux du système de surveillance. C'est pourquoi c'est une « épée à double tranchant » et les décisions de surveillance doivent être affinées. C'est rarement le cas avec les systèmes de surveillance électronique. Les entreprises ne maîtrisent pas toujours les technologies qu'elles acquièrent. Dans le cas de notre étude où l'entreprise avait développé ses propres applications de surveillance, une partie des décisions concernant les paramètres d'application avait été prise par des ingénieurs informaticiens qui n'étaient pas des experts en supervision ou en traitement des appels. Nous voyons également dans ce cas que la conception du système de surveillance électronique aurait dû tenir compte de la différence d'engagement au travail entre un appel entrant ou un appel sortant. Les employés ne réagissent pas seulement à la surveillance en variant leur niveau d'effort pour se livrer à des comportements déviants, une fois que les mesures de performance sont cristallisées dans le système de surveillance électronique, le travail des employés s'appuie sur eux. Ils jouent un rôle de référence dans l'auto-évaluation, un rôle de point focal lorsque des questions de coordination sont en jeu et leur portée définit un espace de contrainte qui encadre l'autodétermination.

Au cours des entretiens, nous avons constaté que tous les conseillers ne réagissent pas de la même manière aux changements apportés par la condition expérimentale. Ils peuvent être satisfaits, stressés, confus ou indifférents. La prise en compte du comportement des employés du centre d'appels dans la vérification des statistiques vient confirmer l'hypothèse selon laquelle les performances peuvent varier en fonction de leur comportement. Certains préfèrent avoir accès à leurs statistiques, d'autres non. De même, les agents préfèrent voir le nombre d'appels en file d'attente pour se préparer, d'autres sont plus sereins sans. Pourquoi ne pas finalement laisser les agents décider librement d'accéder ou non à la file d'attente de l'appelant ? De même, ils

pourraient choisir les indicateurs de performance dont ils ont besoin pour faire leur travail et éliminer ceux qui les dérangent.

2. Les limites de l'expérimentation peuvent restreindre la portée des résultats

Au cours de l'expérience, nous avons fait face à un phénomène d'attrition des réponses aux questionnaires sur la qualité de vie au travail alors que les données de performance fournies par le système d'information de l'entreprise n'ont pas été affectées par ce phénomène. Plusieurs raisons peuvent expliquer cette attrition. Les conseillers devaient répondre tous les quinze jours aux questionnaires et sur toutes les campagnes sur lesquelles ils travaillaient cette quinzaine. Ce qui était une tâche fastidieuse. C'est pourquoi nous avons reçu moins de questionnaires que prévu, environ 30 %. En outre, la période expérimentale a coïncidé avec une période conjoncturelle et un dialogue social difficile pour l'entreprise qui a été négativement ressentie par les conseillers (restrictions budgétaires, risque de liquidations judiciaires…). Cela a pu causer un impact négatif sur le bien-être au travail et l'engagement des conseillers dans l'expérience. Considérant ce contexte, nous serions tentés d'évoquer le mécanisme de la résistance au changement pour expliquer les résultats. Les conseillers téléphoniques n'auraient pas réagi négativement à la diminution de l'intensité de la surveillance électronique, mais pourraient réagir négativement à tout changement introduit. Néanmoins, les principes de l'expérimentation nous ont permis de lutter contre les effets de l'attrition et l'environnement économique puisque les deux groupes ont été impactés de la même manière.

La situation économique tendue pourrait cependant expliquer le comportement des superviseurs. Nous nous attendions à ce qu'ils utilisent le temps gagné à la surveillance dans d'autres tâches telles que le *feedback* ou la formation. Il est possible que, dans ce contexte, ils aient senti plus de pression. À l'instar des enseignants qui, sous une pression plus intense, tendent à exercer davantage de contrôle (Deci et al., 1982), les superviseurs auraient donc utilisé d'autres moyens de surveillance pour atteindre les résultats attendus. Cela suggère un problème concernant la validité de notre condition expérimentale. En effet, nous souhaitions réduire l'intensité de la surveillance alors que les conseillers la jugeaient plus excessive et permanente. L'intensité de la surveillance électronique a été réduite parce que les superviseurs ont objectivement eu accès à moins d'informations. Mais il semble que l'intensité de la surveillance humaine a augmenté. Même si nous n'avons pas donné d'instructions aux superviseurs et que nous ne pouvons pas mesurer leur comportement exact pendant l'expérimentation, nous avons mieux compris les interactions du système de surveillance. Il serait intéressant d'étendre cette expérience dans un environnement conjoncturel plus favorable et en donnant des instructions plus claires aux superviseurs.

3. Une ouverture à un approfondissement du changement d'intensité de la surveillance au travail

Alge et Hansen (2014) indiquent que la recherche sur les effets de la surveillance électronique devrait se concentrer sur l'identification des moyens de rompre la spirale vicieuse d'une surveillance accrue conduisant à la méfiance. Ils pointent l'analyse de la formation des croyances de ceux qui font confiance et de ceux à qui on fait confiance. Nous montrons qu'il est également essentiel de comprendre comment fonctionnent les dispositifs de surveillance de la performance électronique et comment sont liés ces dispositifs à la structuration et à la gouvernance des organisations. Il est possible d'approfondir les pistes de recherche ouvertes par Alge, Ballinger et Green (2004) et Dickinson et Villeval (2008) sur les raisons et la manière dont les organisations et les gestionnaires s'engagent dans la surveillance électronique : quelles sont les représentations de l'effort maximal des évaluateurs (principal) lorsque la surveillance électronique rend les évaluations continues ? comment prennent-ils en compte les stratégies de régulation de l'effort ? Qui décide quel paramètre de conception de la surveillance (intensité, confidentialité, objectifs, etc.) et pour quelles raisons ? Comment le gestionnaire utilise-t-il l'information disponible dans le système pour faire ses jugements en sachant qu'il est limité par des facteurs cognitifs et émotionnels ?

Bates et Holton (1995) recommandent un certain nombre de points à prendre en compte dans la conception d'un système électronique de suivi de la performance : expliciter son objectif, collecter des données sur tous les facteurs susceptibles d'influencer la performance, présenter les données et les dimensions de performance les plus importantes, en rendant l'information facilement accessible par les utilisateurs, en obtenant le consentement des employés et en formant les superviseurs à l'utilisation d'outils et de techniques de surveillance électronique.

D'après Arnaud et Chandon (2013), un autre domaine de recherche future est d'explorer comment les employés construisent leur perception de la surveillance : quels sont les critères et les pratiques qui leur feront sentir que le système de surveillance électronique est attentionné ou coercitif ? Sewell, Barker et Nyberg (2012) proposent un cadre d'analyse pratique qui interroge la légitimité, le but et les règles du droit qui caractérisent les systèmes de surveillance. Des recherches théoriques et empiriques récentes sur la légitimité du contrôle (Schnedler et Vadovic, 2011) et sur l'équité procédurale (Kessler et Leider, 2014) ont commencé à évoluer dans cette direction prometteuse.

Annexe du chapitre 2

4. Questionnaire de mesure de la qualité de vie au travail spécifique à chaque campagne d'appel

Bonjour,
Dans le cadre d'un partenariat de recherche, une expérimentation sociale est implantée au sein de votre établissement. Cette dernière a pour objectif de mettre en lumière des éléments pouvant améliorer votre qualité de vie au travail ainsi que la performance de l'entreprise.
Ce questionnaire a pour objet de recueillir votre opinion et ressenti sur différents aspects de la campagne XXX sur laquelle vous avez travaillé au cours de la semaine écoulée. Merci d'y répondre avant le XX/XX/2014.
Pour répondre aux questions, fiez-vous à votre ressenti. Il n'y a pas de bonnes ou de mauvaises réponses.
Un code anonyme vous sera demandé lors de la première question. Il s'agit du code qui vous a été envoyé lors d'un précédent e-mail. Dans le cadre de l'expérimentation, il est primordial que vous renseigniez cette case avec ce code, et ce pour l'ensemble des questionnaires.
Conformément à la loi, vos réponses à ce questionnaire seront rendues anonymes et confidentielles. Seuls les chercheurs du Centre d'Études de l'Emploi travaillant sur ce projet ont accès à ces données.
Afin que le questionnaire soit valide, merci de répondre à toutes les questions. Sa passation dure une dizaine de minutes.
Merci de votre participation.

Code anonyme : *Obligatoire

1. ACTUALITÉS SUR LA DERNIÈRE QUINZAINE D'EXPÉRIMENTATION

Y a-t-il eu un challenge sur cette campagne ?
Oui
Non

Si oui, y avez-vous participé ?
Oui
Non

Y a-t-il eu une visite du client sur cette campagne ?
Oui
Non

Y a-t-il eu un ou des changements de procédures sur cette campagne ?
Oui
Non

Y a-t-il eu des pannes informatiques sur cette campagne ?
Oui
Non

Si oui, pourriez-vous caractériser ces pannes (liées à la téléphonie, à l'équipement informatique ou au logiciel, interne ou dues au client, travail en mode dégradé ou empêchant tout travail…) ?

Y a-t-il eu d'autres changements sur cette campagne dont vous souhaiteriez nous faire part ?

2. DESCRIPTION DU TRAVAIL

Comment définiriez-vous le travail sur cette campagne ?

	1	2	3	4	
Très difficile					Très facile
	1	2	3	4	
Très déplaisant					Très plaisant
	1	2	3	4	
Très inutile					Très utile
	1	2	3	4	
Très démotivant					Très motivant

Souhaitez-vous apporter des précisions ?

3. SATISFACTION

Dans l'ensemble, êtes-vous satisfait(e) de votre travail sur cette campagne ?

	1	2	3	4	
Pas du tout satisfait(e)					Tout à fait satisfait(e)

Plus spécifiquement, êtes-vous satisfait(e) :

De la prestation en général

	1	2	3	4	
Pas du tout satisfait(e)					Tout à fait satisfait(e)

De l'interaction instaurée avec le client

	1	2	3	4	
Pas du tout satisfait(e)					Tout à fait satisfait(e)

De la qualité du script

	1	2	3	4	
Pas du tout satisfait(e)					Tout à fait satisfait(e)

De la qualité du fichier client

	1	2	3	4	
Pas du tout satisfait(e)					Tout à fait satisfait(e)

Du soutien et de l'aide apportés par les conseillers travaillant également sur cette campagne

	1	2	3	4	
Pas du tout satisfait(e)					Tout à fait satisfait(e)

Du soutien et de l'aide apportés par le/la superviseur ayant la responsabilité de cette campagne

	1	2	3	4	
Pas du tout satisfait(e)					Tout à fait satisfait(e)

De la reconnaissance de votre travail sur cette campagne par votre hiérarchie

	1	2	3	4	
Pas du tout satisfait(e)					Tout à fait satisfait(e)

De la formation reçue sur cette campagne

	1	2	3	4

Pas du tout satisfait(e)	Tout à fait satisfait(e)

De l'amplitude de vos horaires de travail sur cette campagne

1	2	3	4
Pas du tout satisfait(e)			Tout à fait satisfait(e)

De la manière dont cette campagne contribue à votre prime

1	2	3	4
Pas du tout satisfait(e)			Tout à fait satisfait(e)

4. FACTEURS DE STRESS ET DE FATIGUE

Dans l'ensemble, estimez-vous que votre travail sur cette campagne est fatigant ?

1	2	3	4	5	6	7
Pas du tout fatigant						Extrêmement fatigant

Dans l'ensemble, estimez-vous que votre travail sur cette campagne est stressant ?

1	2	3	4	5	6	7
Pas du tout stressant						Extrêmement stressant

Facteur de stress et de fatigue : *Obligatoire
Citez le facteur de stress et de fatigue le plus important (charge de travail, marges de manœuvre, manque de soutien social, etc.) que vous avez rencontré lors de la réalisation de cette campagne, au cours de la semaine écoulée.

Concernant ce facteur qui vous stresse et vous fatigue dans le cadre de cette campagne, diriez-vous qu'il...

1	2	3	4	5	6	7
Est contrôlable						Est incontrôlable

1	2	3	4	5	6	7
Est plutôt lié à vous-même						Est plutôt lié à la situation

Y a-t-il eu un épisode ou un moment particulièrement stressant lors de cette dernière quinzaine, pourriez-vous le décrire ?

5. UTILISATION DES STATISTIQUES PAR LES SUPERVISEURS

Pour chaque affirmation, veuillez cocher le chiffre correspondant le plus à votre opinion de 1 : pas du tout d'accord à 4 : tout à fait d'accord. Les superviseurs se servent des statistiques, dans le cadre de cette campagne :

Pour améliorer la productivité de l'entreprise (rendre des comptes aux clients donneurs d'ordres, améliorer la répartition des conseillers sur les campagnes…)

1	2	3	4
Pas du tout d'accord			Tout à fait d'accord

Pour m'aider à progresser dans mon travail

1	2	3	4
Pas du tout d'accord			Tout à fait d'accord

Pour contrôler mon travail

1	2	3	4

Pas du tout d'accord	Tout à fait d'accord

Pour s'assurer que je ne rencontre pas de difficultés dans mon travail (situation tendue avec un client, problème technique...)

1 2 3 4

Pas du tout d'accord	Tout à fait d'accord

Souhaitez-vous apporter des précisions ?

6. ENVIRONNEMENT DE TRAVAIL

Pour chaque affirmation, veuillez cocher le chiffre correspondant le plus à votre opinion de 1 : pas du tout d'accord à 4 : tout à fait d'accord.

Je trouve la surveillance informatique excessive

1 2 3 4

Pas du tout d'accord	Tout à fait d'accord

Je n'ai pas le sentiment d'être surveillé(e) en permanence

1 2 3 4

Pas du tout d'accord	Tout à fait d'accord

Je me sens libre de réaliser mon travail comme je l'entends

1 2 3 4

Pas du tout d'accord	Tout à fait d'accord

Le superviseur en charge de cette campagne me fait confiance pour bien réaliser le travail

1 2 3 4

Pas du tout d'accord	Tout à fait d'accord

Souhaitez-vous apporter des précisions ?

7. POUR TERMINER…

Souhaitez-vous ajouter des remarques (positives ou négatives) ou des suggestions ?

Chapitre 3.
Négocier les règles d'évaluation pour développer l'accompagnement à la réussite des étudiants

I. Introduction

En 2016, l'indicateur de réussite à l'université, calculé par le Ministère de l'Enseignement Supérieur, mesurait que parmi les néo-entrants à l'université inscrits en 2012-2013, trois étudiant sur dix décrochaient leur licence en trois ans (Maetz, 2017). En tenant compte des trajectoires de réorientation ou de redoublement, seuls quatre étudiants sur dix de cette même promotion ont obtenu leur licence en quatre années. Depuis l'entrée en vigueur de l'arrêté du 23 avril 2002, relatif aux études universitaires conduisant au grade de licence, cet indicateur est stable et il ne semble pas connaître d'inflexion à la hausse (voir Annexe n°1).

Ce faible nombre de diplomations en trois ans conduit à définir l'échec des étudiants à l'université comme l'ensemble des situations de non-réinscription, de réorientation ou de redoublement. Ainsi, chaque année, 30 % des effectifs ne se réinscrivent pas à l'université, sans qu'il soit possible de déterminer, à un niveau national, l'issue de leur sortie : vers un autre établissement de l'enseignement supérieur ? ; vers l'emploi ? ; vers l'inactivité ?

Malgré cette catégorisation administrative, l'échec des étudiants ne bénéficie pas encore d'une définition conceptuelle (Millet, 2012). Ce flou théorique s'explique d'abord par le fait que l'échec recouvre des stratégies individuelles différentes. Par exemple, un étudiant qui n'aurait pas décroché de formation sélective de son choix serait en position d'attente sur les bancs de l'université pour obtenir ultérieurement un cursus préféré, ce qui augmenterait le taux de redoublement ou de réorientation. L'abandon peut également se traduire par une sortie vers l'emploi.

Il convient également de prendre en compte des phénomènes sociaux qui conditionnent le choix des étudiants à poursuivre leur parcours d'études supérieures ou à l'abandonner. Par exemple, les informations sur la nature et le contenu disciplinaire des enseignements, sur les prérequis en termes de compétences et de socle de connaissances, enregistrées dans le système d'orientation des bacheliers sont des vecteurs qui conditionnent durablement les affectations des étudiants.

Enfin, l'investissement dans les cursus de formation est un élément qui facilite ou détériore l'accueil et les conditions d'étude des étudiants.

1. Les inégalités sociales face à l'échec à l'université et les outils de remédiation

L'échec est un phénomène qui attire l'attention et qui a conduit les chercheurs à étudier les caractéristiques sociodémographiques de la population d'étudiants qui sont en échec à l'université (pour une revue de la littérature récente, se reporter à Aina et al., 2018 et Duguet et al., 2016). Il ressort de ces études que la population étudiante est inégalement atteinte par l'échec à l'université.

Le genre est un premier facteur de la réussite : les hommes sont ainsi moins susceptibles que les femmes d'accomplir un parcours de licence en trois ans. Il existe également une hiérarchie de la réussite selon le type de baccalauréat obtenu avant d'entrer à l'université : les diplômés du baccalauréat technologique réussissent moins bien que ceux diplômés du baccalauréat général, tandis que les diplômés du baccalauréat professionnel sont ceux qui ont le moins de chance de réussir dans un parcours linéaire. Le travail salarié des étudiants est une autre source d'échec : travailler plus de 20 heures par semaine (Beffy et al., 2009) et pendant des périodes supérieures à 6 mois (Dmitrijeva et al., 2014) se traduit par davantage d'insuccès aux examens. Enfin, ce serait l'ensemble du parcours antérieur de l'étudiant qui expliquerait le plus l'échec : le fait d'avoir effectué un redoublement, le fait de ne pas avoir eu de mention et le fait d'entrer avec un âge plus élevé que les partenaires de promotion sont autant de facteurs qui pénalisent les étudiants dans leurs parcours d'études.

Cette diversité des profils d'échec s'explique d'abord par l'élargissement du spectre du recrutement des étudiants dans les établissements de l'enseignement supérieur. En effet, l'ensemble des universités françaises a fait face à une importante massification des effectifs étudiants. Cette massification est consécutive à la « seconde démocratisation de la scolarisation », c'est-à-dire l'ouverture des établissements de l'enseignement supérieur à un nouveau public qui n'avait pas eu jusqu'alors la possibilité de suivre des cursus longs de formation.
Cette ouverture est la conséquence conjointe de la succession de politiques éducatives volontaristes[17] et de la mesure des bienfaits de l'investissement dans le capital humain avant d'entrer sur le marché du travail (Becker, 1962). Alors que dans l'année 1960-1961, 309 700 personnes étaient inscrites dans l'enseignement supérieur, cinquante ans plus tard c'est sept fois plus d'étudiants qui sont entrés dans des formations post-baccalauréat. Pendant cette période, ce sont les universités qui ont été les plus importantes contributrices à la massification du public étudiant, en accueillant chaque année entre 60 % et 70 % de l'effectif des inscrits.

La conséquence de cette démocratisation de l'enseignement supérieur est que les établissements ont davantage accueilli des étudiants de première génération, c'est-à-dire des étudiants dont les parents n'avaient pas suivi de cursus de l'enseignement supérieur. De plus, l'origine sociale des étudiants s'est davantage diversifiée, puisque les enfants de parents des catégories

[17] Nous citerons les plus importantes parmi celles qui ont jalonné les cinquante dernières années : la loi n°68-978 du 12 novembre 1968 sur l'enseignement supérieur qui a réformé administrativement l'université en lui conférant davantage d'autonomie dans son organisation et sa gouvernance, qui a encouragé le développement de la pluridisciplinarité et qui a créé le statut d'enseignant-chercheur ; la loi n°75-620 du 12 juillet 1975 relative à l'éducation qui instaure la mise en place du « collège unique » en y ouvrant l'accès à tous les élèves gratuitement et en y incorporant une homogénéisation des disciplines enseignées ; la loi n°89-486 du 10 juillet 1989 qui a fixé l'objectif de 80 % d'une classe devant être titulaire d'un baccalauréat. Ces réformes ont permis de configurer le système éducatif français vers une structure de formation longue des individus avant leur entrée sur le marché du travail.

socioprofessionnelles employées et ouvrières ont davantage franchi les portes de l'université. Le public accueilli est donc devenu plus hétérogène.

Alors que le taux d'échec était considéré aussi bien comme une sélection qui s'opérait naturellement que comme un gage d'excellence du diplôme délivré et un signe de l'échec personnel des étudiants à réussir à l'université, la massification a redéfini son interprétation. Ce phénomène pose désormais la question de la qualité de l'organisation de la formation des étudiants des universités parce qu'il traduirait une situation particulière : la difficulté pour les universités d'accompagner les flux d'étudiants dans des conditions de réussite plus satisfaisantes. Les universités paraissent ainsi accuser un retard dans l'adaptation des formations à ces « nouveaux publics ». Cette hétérogénéisation ne s'accompagnerait d'un encadrement plus spécifique qui tiendrait compte des caractéristiques du nouveau public étudiant.

Nous pourrions néanmoins nuancer cet argumentaire en présentant la multitude de dispositifs de remédiation à l'œuvre à l'université, tant au niveau national qu'au niveau local. Le « Plan Réussite en Licence », doté de 730 millions d'euros sur la période 2007-2012 a reformulé la licence pour promouvoir une spécialisation progressive sur les trois années de la préparation du diplôme et un renforcement de l'encadrement pédagogique grâce à davantage d'heures hebdomadaires pour accompagner les étudiants. En 2013, le « Plan National Vie Étudiante » a mis l'accent sur les conditions de vie des étudiants en développant l'accès au logement, à la santé et en réduisant les démarches administratives. Enfin, pour la période 2017-2022, le « Plan Étudiants » prévoit une meilleure lisibilité des formations ainsi qu'un approfondissement de l'orientation et de l'accès des lycéens à l'enseignement supérieur.
Au niveau local, le développement du tutorat qui vise à apporter une forme d'aide individualisée en dehors du contexte de la classe, la création de composantes de formation à la méthodologie pour favoriser le développement de compétences complémentaires à celles acquises pendant la formation et l'investissement dans des modules de renforcement des acquis liés à la discipline étudiée sont autant d'innovations qui ont tenté de favoriser l'accueil des étudiants[18].

[18] Dans le contexte français, il existe peu d'évaluations d'impact de ces dispositifs sur l'échec à l'université. À notre connaissance, une seule étude a estimé l'incidence du « Plan Réussite en Licence » sur les étudiants de l'Université de Bourgogne (Morlaix et Perret, 2013). Tout en relevant les difficultés de mesures de l'effet causal des dispositifs d'accompagnement, à cause de leur nombre élevé sur le terrain et de la méthodologie employée, les auteures ont mesuré un effet nul ou non significatif des dispositifs. Les chercheurs de l'Institut de Recherche sur l'Éducation (IREDU, EA7318) notent que c'est finalement la diversité des opérations qui conduisent à brouiller la mesure des effets sans pouvoir attribuer de lien de causalité propre à un dispositif particulier (accueil et suivi pédagogique renforcé, méthodologie de travail universitaire, réorientation des étudiants en situation d'échec…).
Les rares recherches sur les dispositifs d'accompagnement méthodologiques ou pédagogiques montrent que le plus souvent ils sont facultatifs et touchent rarement ceux qui en auraient le plus besoin (Morlaix et Lambert-Le Mener, 2015).

2. Les limites de l'évaluation des connaissances comme facteur explicatif de l'échec à l'université

D'un point de vue organisationnel, la reconfiguration de l'université pour tenir compte de la diversité du public étudiant ne semble pas suffire pour améliorer le parcours de formation des étudiants. Il nous semble donc opportun de mobiliser la relation enseignant-étudiant dans les établissements de l'enseignement supérieur (Correa et Gruver, 1987) pour rendre compte des tensions qui existent dans les parcours d'apprentissage et dans les modalités d'enseignement spécifique à l'université. Sous plusieurs hypothèses, cette relation peut se caractériser comme une relation d'agence particulière (Zubrikas, 2016) au sein de laquelle les pratiques d'évaluation des connaissances jouent le rôle du dispositif de vérification de l'information. Dans cette relation, il est essentiel de (re)définir le rôle et les comportements des étudiants et des enseignants ainsi que la place du dispositif d'évaluation des connaissances.

D'une part, les aspects motivationnels des étudiants, c'est-à-dire leurs stratégies de progression dans les études supérieures, seraient des éléments majeurs dans la décision d'abandonner les études. En économie de l'éducation, un champ de recherche s'est développé autour de la notion d'effort d'apprentissage des étudiants, comparable à l'idée d'effort au travail dans le cadre de la relation d'emploi (Bonesrønning et Opstad, 2012). La motivation à l'apprentissage est un état dynamique qui a ses origines dans les perceptions qu'un élève a de lui-même et de son environnement et qui l'incite à choisir une activité, à s'y engager et à persévérer dans son accomplissement afin d'atteindre un but (Viau, 1998). Ainsi, la motivation joue un rôle positif et significatif sur la moyenne annuelle obtenue par les étudiants, tandis que l'amotivation exerce un impact négatif sur les résultats aux examens (Duguet, 2014 ; Lambert-Le Mener, 2012).

D'autre part, dans l'essentiel des dispositifs que nous avons décrit, rien n'est prescrit sur la conduite de la formation des étudiants. Pourtant, l'acteur majeur de l'enseignement à l'université est bien le formateur, qu'il soit doctorant, vacataire, professeur agrégé du secondaire ou enseignant-chercheur. Sa position est donc remise en question.

Les pratiques des enseignants font l'objet d'une analyse intensive par les chercheurs en science de l'éducation (Adangnikou, 2008 ; Duguet et Morlaix, 2012). Bertrand (2014) indique que la pédagogie constitue un facteur déterminant pour la réussite des étudiants. Il considère en ce sens qu'il est important de faire évoluer les pratiques d'enseignement et d'apprentissage, notamment dans le but de pouvoir faire face aux différents publics étudiants. Duguet (2014) a mis au jour le fait que trois dimensions des pratiques pédagogiques liées aux interactions avec les étudiants, à la façon d'organiser le cours et à l'attitude des enseignants, exercent un réel effet significatif, toutes choses égales par ailleurs, sur la réussite.

Cependant, hormis cette analyse, Michaut (2012) énonce que les effets des pratiques enseignantes sur la réussite restent aujourd'hui une « zone d'ombre », personne ne s'étant jusqu'à présent interrogé sur « l'effet enseignant » ou sur les pratiques d'évaluation des connaissances.

Dans cette relation, l'évaluation est un dispositif que les enseignants utilisent pour collecter des preuves du niveau d'apprentissage atteint par les étudiants. Il permet de mesurer ce que les étudiants connaissent, comprennent et sont capables de réaliser afin de produire une inférence sur l'apprentissage. C'est un outil de jugement sur les performances des étudiants, sur leurs forces et leurs faiblesses et de contrôle de leurs progrès par rapport à un but fixé. Sa finalité est d'aider à la notation et à la certification du niveau atteint par les étudiants pour la délivrance du diplôme (Bloom et al. 1971 ; McMillan, 2015 ; OCDE, 2013).
L'évaluation est également un dispositif qui permet d'inciter à l'apprentissage des étudiants en donnant les informations sur les attentes de l'enseignant ou de l'institution, en termes d'acquisition de niveau de connaissances, de compétences et de savoir-faire. C'est un outil qui conditionne l'apprentissage et la motivation à apprendre en développant les compétences de planification et d'autonomie (Gibbs et Simpson, 2002). Il existerait conceptuellement cinq dimensions qui permettent de mesurer la qualité de l'évaluation (OCDE, 2013). Le critère de validité repose sur l'idée selon laquelle ce sont les aspects pertinents de la performance qui doivent être mesurés. Le critère de fiabilité permet de créer des évaluations qui ne sont pas influencées par des considérations subjectives. Le critère de transparence rend les règles de l'évaluation claires et permet la consultation externe des résultats. Le critère de facilité d'utilisation assure que les procédures d'évaluation sont compréhensibles par l'évalué et par d'autres parties prenantes (l'équipe enseignante, l'administration, les parents, etc.). Enfin, selon le critère d'équité, l'évaluation ne doit pas favoriser des étudiants parmi d'autres, par exemple lorsque les tests introduisent des biais linguistiques qui pénalisent des étudiants qui ne comprendraient pas les consignes.

Dans les universités, l'évaluation des connaissances repose sur deux modalités définies au niveau national. La première concerne l'examen terminal de fin de semestre, le partiel, qui est destiné à valider l'appropriation complète des connaissances d'un module de formation afin d'acquérir un nombre correspondant de crédits dans le cadre du système européen de crédits transférables (« *l'european credits transfer system* » - ECTS). Ces derniers conditionnent l'obtention du diplôme. La seconde consiste à évaluer de façon continue l'apprentissage à travers plusieurs épreuves au cours d'un semestre.
Au niveau local, les universités sont relativement autonomes pour l'application de ces règles dans la forme (oral ou écrit), le contenu (exercice, questions ouvertes, fermées, à réponses multiples,

dossier de recherche documentaire), le degré d'individualisation, la durée, la temporalité, et le nombre des examens à faire passer. Ce sont les commissions de la formation et de la vie universitaire (CFVU) qui ont la prérogative de valider les modalités générales et spécifiques des contrôles de connaissances dans les offres de formation. Ces dernières sont préalablement validées par les conseils d'unité de formation et de recherche (UFR).

En définitive, la création d'un examen est à la discrétion de l'enseignant, qui est le seul concepteur et créateur des critères, des règles et des pondérations de l'évaluation. Ce rôle primordial lui confère une autorité et un pouvoir de contrôle sur l'apprentissage des étudiants. Mais cette dimension de la profession entre en conflit avec d'autres fonctions qui lui incombent : la recherche, l'enseignement, la responsabilité de formations, les activités professionnelles en parallèle de la formation. Il existe donc un arbitrage dans le temps dévolu à l'évaluation.

Cet arbitrage conduit à choisir des modalités d'évaluation les plus rapides possible, qui demandent un temps de jugement le plus court possible, afin d'attribuer du temps à des tâches davantage valorisées. En d'autres termes, le coût d'opportunité de l'évaluation est élevé. Paradoxalement, la valeur attribuée à cette tâche est négativement reliée au temps nécessaire pour réaliser l'évaluation. En effet, à l'évaluation s'ajoute la période de correction et le nombre d'étudiants à évaluer augmente tendanciellement.

De plus, contrairement à l'enseignement primaire et secondaire qui est prescrit par des programmes nationaux, la spécificité de la connaissance transmise à l'université repose sur des champs disciplinaires en perpétuelle évolution. Cette particularité se répercute sur les modes d'évaluation qui dans le premier cas peuvent être standardisés et répliqués dans tous les établissements de formation d'un territoire donné, tandis que dans le second, l'évaluation dépend des pratiques disciplinaires, des caractéristiques des enseignants et de l'appréciation générale sur ce qui doit être appris et compris.

Par conséquent, les formes d'évaluation à l'œuvre à l'université sont davantage de nature à opérer un classement des étudiants afin de pouvoir sélectionner ceux qui seront les plus aptes à continuer leur formation. Cette utilisation de l'évaluation a pour finalité de contrôler l'apprentissage, ce que les chercheurs spécialistes de l'évaluation scolaire ont dénommé comme étant l'évaluation « *summative* ». Elle correspond aux évaluations générales standardisées se situant à chaque fin de cycle d'enseignement qui permettent de délivrer un diplôme nécessaire pour passer à un niveau supérieur de formation. Par exemple, le baccalauréat correspond à cette définition.

Cette manière d'évaluer permet de répondre aux attentes de l'institution d'enseignement supérieur au premier rang desquelles se situe la capacité de certifier le niveau des étudiants en formation. Cependant, l'évaluation « *summative* » produirait des conséquences néfastes sur le

comportement des étudiants, un effet « *backwash* », c'est-à-dire qu'elle créerait un phénomène contraire à ce qu'elle est censée produire (OCDE, 2013).

Elle développe chez les étudiants une culture du résultat, afin de réussir à passer les épreuves d'examen de connaissances. Cette culture se traduit par le développement de stratégies d'optimisation de la réussite, pointées dans un rapport de l'Inspection générale de l'administration de l'éducation nationale et de la recherche (Gauthier et al., 2007). Ces stratégies passent par le choix d'enseignements plus favorables à la validation d'un semestre, au développement de transmission de cours clés en main, au choix d'un module où l'enseignant pratique une notation plus favorable.

Cette réaction à l'organisation traditionnelle des examens est conditionnée par la peur de l'insuccès : en effet, les étudiants ne souhaitent pas remplir les statistiques de taux d'échec. La conséquence directe de la culture du résultat, réaction à l'existence de l'évaluation « *summative* », est la génération de stress et d'anxiété pour les étudiants qui doivent choisir la stratégie la plus adaptée pour réussir (Sambell et al., 2013). Il s'agit d'une réaction identique à celle qui est constatée pour l'évaluation des performances dans le cadre de l'entreprise (voir chapitre 1).

Les effets négatifs de cette culture du résultat pourraient se borner à conditionner un état émotionnel plus ou moins favorable à la réussite qui correspondrait à un certain type d'étudiant si elle ne contenait pas en germe une dimension plus dommageable à la réussite. En effet, elle provoque également une distanciation des étudiants avec le développement d'un goût, d'une appétence ou d'une motivation intrinsèque pour l'apprentissage, qui renforcerait la curiosité et l'esprit critique sur l'environnement de l'étudiant. Par conséquent, elle provoque une dépossession de l'apprentissage, c'est-à-dire que les étudiants n'apprennent pas pour l'acquisition et la conservation durable des connaissances et des compétences, mais uniquement pour réussir à un examen. L'appellation que les étudiants donnent à ce fonctionnement consistant à se soumettre à une succession « d'enseignements », avec les contrôles associés, en se dépêchant de passer au suivant, et de jeter jusqu'aux notes et supports des cours : c'est le « coup de la chasse d'eau » (Gauthier et al., 2007).

En somme, l'évaluation des connaissances entre directement en conflit avec les missions de l'enseignement supérieur, à savoir le développement de l'apprentissage tout au long de la vie et la capacité à être autonome pour résoudre des problèmes inconnus (Birenbaum, 1996). Elle empêche la capacité à avoir confiance en soi et en ses compétences.

In fine, l'évaluation « *summative* » provoquerait un processus qui s'autoalimente : la peur de se retrouver dans les statistiques de l'échec à l'université déclenche des stratégies qui entrent en conflit avec ce qui serait la meilleure stratégie pour contrer l'échec. Alors que l'évaluation traditionnelle de l'université devrait permettre de valider l'acquisition des compétences spécifiques

à une économie de la connaissance, elle renforce d'autres compétences qui finalement produisent de l'échec : les étudiants ne développent pas les compétences attendues et échouent aux examens dédiés à valider l'acquisition de l'autonomie, de l'adaptation, de la prise d'initiative, etc.

3. Les marges de manœuvre possibles pour améliorer le dispositif d'évaluation des connaissances

Malgré ces conséquences négatives, l'évaluation resterait résistante au changement et ce serait le dispositif le plus conservateur de l'université (Bloxham, 2016). Pourtant, une seconde forme d'évaluation a été conceptualisée pour améliorer les pratiques et dépasser ces effets négatifs. L'évaluation « formative » est ainsi dédiée à favoriser l'apprentissage (Scriven, 1967 ; Brown et Knight 1994, Black et William, 1998). L'évaluation pour l'apprentissage modifie le rapport de l'administration de l'enseignement supérieur, des enseignants et des étudiants à l'évaluation. L'idée étant qu'une nouvelle culture de l'évaluation contribue à améliorer les résultats des étudiants à travers des pratiques d'apprentissage et d'enseignement perfectionnées.

Dans ce cadre, l'évaluation pour l'apprentissage intègre le retour sur le résultat, c'est-à-dire la production d'un commentaire sur le résultat à un examen de connaissances pour faire comprendre où se situent les erreurs et comment y remédier, commentaire qui se conclut par un nouvel examen pour apprécier les pistes d'amélioration. Le « *feedback* » est un des rares concepts généraux en science de l'éducation clairement validés par les chercheurs (McKeachie, 1962). Derrière cette idée, l'évaluation doit servir à informer de façon pertinente les apprenants aussi bien qu'elle permet d'informer les autres parties prenantes des résultats de l'évaluation.

L'évaluation « formative » favorise la transparence de l'évaluation. Elle intègre l'explicitation précise de l'articulation du critère d'évaluation et les objectifs à atteindre. Un lien clair et explicite entre ce qui est enseigné et ce qui est évalué permet aux étudiants de se focaliser sur des points importants du cours. Cela permet de s'assurer que la charge de travail est réaliste et que l'évaluation n'est pas menaçante. Par conséquent, cela renforce l'autorégulation de son propre travail, c'est-à-dire qu'elle rend capable de se situer soi-même dans l'acquisition des connaissances. Le concept de littéracie de l'évaluation des étudiants traduit ce phénomène (Rust et al., 2003). Cela permet d'augmenter leurs compétences d'évaluation des pairs et d'auto-évaluation. *In fine*, l'évaluation pour l'apprentissage permet également de former les étudiants à maîtriser la capacité d'évaluer des preuves, de juger des situations et les circonstances, de tirer des conclusions et d'agir en conséquence.

La liberté est ainsi donnée à la prise de risque et à la possibilité de faire des erreurs. Elle permet de diagnostiquer les besoins en apprentissage et elle renforce l'engagement des étudiants dans le processus d'apprentissage. Le curseur est davantage mis sur la confiance envers l'apprenant dans la construction de son savoir plutôt que sur une relation asymétrique où le savoir est détenu et transmis unilatéralement.

Cette manière d'évaluer est largement diffusée dans le cycle primaire et secondaire de l'enseignement. Black et William (1998) ont ainsi conclu sans équivoque qu'après vingt années de recherche sur les pratiques d'évaluation pour l'apprentissage, ces dernières améliorent la réussite des élèves. En revanche, ce glissement sémantique peine à s'inscrire dans les pratiques de l'enseignement supérieur (Boud et Falchikov, 2007). Pourtant, il permet de renforcer l'idée de l'acquisition d'une culture de l'apprentissage tout au long de la vie et d'un répertoire propre à l'évaluation à appliquer dans le cadre du travail dans une économie de la connaissance.

4. Problématique de la recherche

En résumé, dans le cadre de la relation enseignant-étudiant, le dispositif de l'évaluation des connaissances et des compétences tel qu'il se réalise à l'université peut sembler mal calibré à la fois pour révéler l'information cachée sur l'apprentissage des étudiants et pour les inciter à un apprentissage durable. Il parait nécessaire de trouver un nouvel équilibre de l'évaluation entre sa dimension de mesure de performance et celle de promotion de l'apprentissage. En terme organisationnel, il s'agit de développer un nouveau contrat optimal d'évaluation des connaissances dans la relation étudiant-enseignant de l'enseignement supérieur.

Nous faisons l'hypothèse que la conception de l'évaluation comme quelque chose qui implique les deux parties de la relation, ouverte à la négociation, qui mette l'accent aussi bien sur le processus que sur le résultat, doit permettre de renforcer et de développer des compétences pouvant conduire les étudiants à améliorer leurs résultats à l'université. L'étudiant doit participer au processus et il doit être capable d'auto-évaluer son travail pour lui donner du sens. Cela impliquerait une rupture dans la balance du pouvoir. Les étudiants auraient plus de contrôle, de choix et de flexibilité quant à la formation et sur le développement de leur propre apprentissage, mais également sur les modalités de leur évaluation.

Dans ce contexte, nous souhaitons déterminer dans quelle mesure une relation enseignant-étudiant dans laquelle réside davantage de réciprocité, où chacun a un rôle et une voix pour influencer la participation dans l'évaluation permettrait-elle de favoriser la réussite des étudiants ?

Nous avons tenté de répondre à cette question en expérimentant une forme nouvelle d'évaluation qui passe par la négociation et la co-construction des critères de notation de la participation des étudiants au sein de deux modules de formation de la licence d'Économie Gestion de l'Université Paris-Est Marne-la-Vallée. Cette expérimentation exploratoire a permis de faire dialoguer un chargé de travaux dirigés avec des groupes d'étudiants sur ce qu'il est pertinent de considérer quand il s'agit d'examiner et donner un score au niveau de la participation dans la formation. Nous mesurerons l'impact de ce dispositif sur leurs résultats aux examens de fin de semestre ainsi que sur la probabilité de valider le semestre.

Dans un premier temps, nous réaliserons une revue de la littérature de l'économie de l'éducation pour montrer qu'il existe des outils conceptuels qui permettent de justifier de l'utilisation de la négociation de l'évaluation dans la relation enseignant-étudiant, ainsi que des études en sciences de l'éducation qui ont commencé à produire des résultats novateurs sur le partenariat dans l'évaluation. Dans un deuxième temps, nous présenterons le protocole expérimental et le recueil de données. Enfin, les résultats de cette étude montreront que la participation des étudiants à la construction des modalités d'évaluation joue un rôle substantiel sur la réussite des étudiants.

II. Revue de la littérature

La théorie du capital humain procure un cadre unifié des décisions des étudiants à investir dans l'éducation, la durée et le type de scolarité et de formation et indique comment ces investissements déterminent leur productivité et leur revenu. De nombreuses recherches s'inscrivent dans ce cadre théorique au sujet des incitations à rester à l'école, sur le retour sur investissement dans des années de scolarité supplémentaires, sur la qualité des formations, sur le choix de la formation professionnelle, sur les variations du salaire, et sur les liens avec le marché du travail (Card, 1999).

En revanche, ce n'est pas ce cadre théorique qui a été employé pour l'analyse économique des décisions prises au sein des écoles et des classes, c'est-à-dire le processus d'enseignement et d'éducation. À l'instar de l'analyse des décisions individuelles prises par les ménages ou les firmes, il s'agit d'examiner l'organisation générale et l'allocation des ressources au sein des structures d'enseignement (plutôt que la structure des leçons ou la programmation des enseignements) (Brown et Saks, 1981 ; Brown et Saks, 1987).

1. La modélisation de la relation enseignant-étudiant

Dans cette « microéconomie de l'instruction », les questions correspondent à celles qui se posent traditionnellement pour analyser les choix individuels : « qui choisit ? », permet de distinguer

l'individu sur lequel portera l'analyse du comportement, qu'il s'agisse de l'enseignant, de l'étudiant, des parents, des responsables de formation ; « qu'est-ce qu'il souhaite accomplir ? » correspond à l'objectif de chaque individu dans le cadre de l'instruction, par exemple celui de contribuer à procurer un niveau d'apprentissage, ou de recevoir un revenu maximal ; « comment est-il capable de prendre ses décisions ? » décrit les contraintes et l'environnement qui influencent le choix, par exemple le temps passé dans les activités d'apprentissage, les ressources physiques de l'enseignement, le comportement des autres acteurs.

Nous décrirons dans un premier temps le cadre théorique de la microéconomie de l'instruction et la principale contribution de Correa et Gruver (1987) pour modéliser les interactions entre l'enseignant et l'étudiant. Dans ce cadre, nous présenterons la possibilité de négocier les termes de l'enseignement.

Nous détaillerons ensuite une vision alternative de l'enseignement, qui tient davantage compte de la spécificité de l'évaluation, dans un contexte d'aléa moral. Enfin, nous présenterons la possibilité d'inclure la négociation dans cette dimension.

a. La modélisation canonique considère l'enseignement comme un bien public

L'enseignement ou l'instruction (« *schooling* ») correspond à un processus dans lequel les préférences de l'étudiant et les préférences de l'enseignant sont associées avec d'autres ressources pour former l'apprentissage (Brown et Saks, 1981). En tenant compte des apports de la psychologie, les économistes ont appelé cette relation « la courbe de l'apprentissage ». Elle équivaut formellement à la fonction de production de la firme, parce qu'elle correspond à l'idée qu'il faut associer des intrants, c'est-à-dire des facteurs d'apprentissage, qui sont transformés d'une certaine manière pour produire un bien, en l'occurrence le résultat de l'apprentissage.

Considérons une certaine courbe d'apprentissage pour un étudiant donné et dans une formation donnée,

$$y = f(x_1, \dots, x_n)$$

où la variable y correspond au résultat de l'apprentissage, l'ensemble $\{x_1, \dots, x_n\}$ désigne la quantité de facteurs de production de l'apprentissage, et la fonction $f(.)$ est une unique transformation de ces facteurs.

Le résultat de l'apprentissage peut être considéré comme un indicateur de la réussite des étudiants, ou au minimum l'indicateur du niveau de l'étudiant après être passé dans l'instruction. De façon pragmatique, il correspond à la note obtenue à un examen après un temps passé dans le cursus de formation. Dans ce modèle, la réussite liée à l'apprentissage est assimilable à un bien

public et dispose de toutes ses caractéristiques : non-rivalité dans la consommation (ce qu'en consomme une personne n'en diminue pas la consommation des autres) ; impossibilité d'exclusion (on ne peut empêcher personne de bénéficier de ce bien) ; impossibilité d'abstention (personne ne peut s'exclure du bénéfice de ce bien). Il en résulte que le résultat de l'apprentissage ne peut pas être évalué à un prix de marché et par conséquent il ne peut pas être financé à partir du produit de sa vente. De plus, les mécanismes de marché ne sont pas capables d'en générer un montant convenable parce qu'il engendre des externalités dont ces derniers ne peuvent pas tenir compte.

Les facteurs de production peuvent être divisés en deux catégories. La première correspond à la part des caractéristiques individuelles de l'étudiant qui ne sont pas sujettes au changement. Traditionnellement, cela correspond aux aptitudes, c'est-à-dire à la disposition innée d'une personne à réaliser une activité, en l'occurrence les savoir-faire et les savoir-être acquis en dehors de la formation, grâce à l'entourage familial par exemple. Cela inclut également le niveau atteint antérieurement à une formation donnée.

La deuxième inclut un ensemble de facteurs qui peuvent varier à l'intérieur du système éducationnel : ce sont les variables de choix d'apprentissage. Elles peuvent se répartir entre les deux acteurs de la relation d'apprentissage : les décisions des enseignants sur leur façon d'instruire leurs étudiants, le degré de pédagogie ou de discipline, sur leur manière d'évaluer le niveau d'apprentissage ; les décisions des étudiants sur l'effort qu'ils consacrent à l'apprentissage.

Il serait également possible d'ajouter une troisième catégorie de facteurs de production qui caractériseraient l'environnement de la relation. Il s'agit des décisions du responsable de formation, par exemple sur la taille des classes ou le matériel éducatif, qui sont un critère de qualité de la formation. Il s'agit aussi du rôle que joue le comportement des camarades de classe sur le niveau d'apprentissage, que l'on nomme les effets de pairs. Ces derniers interagissent avec les caractéristiques de l'enseignant, par exemple lorsqu'il s'agit pour lui de demander plus ou moins de discipline dans la classe ainsi qu'avec les caractéristiques des étudiants, lorsqu'ils bénéficient de l'aide d'un camarade pour la compréhension d'un exercice. Dans la suite de l'exposé, nous ne tiendrons pas compte de ce troisième facteur afin de ne pas complexifier l'analyse.

D'un point de vue théorique, la forme fonctionnelle de la courbe d'apprentissage revêt une importance particulière pour apprécier le pouvoir prédictif des différents changements de facteurs de production sur le résultat final de l'apprentissage, c'est-à-dire pour être en mesure de capturer l'effet marginal d'un facteur donné tout en gardant les autres constants. Un individu qui chercherait à maximiser la valeur totale de l'apprentissage nette de ses coûts utilisera de plus en plus de facteurs l'influençant à mesure que le résultat additionnel est supérieur au coût

d'acquisition d'une unité supplémentaire de facteur de production. Par exemple, un enseignant qui cherche à augmenter le résultat de l'apprentissage, en utilisant une heure supplémentaire de lecture d'ouvrages ou une heure supplémentaire d'exercices, arbitrera sur celle qui procure le plus de gain marginal sur les résultats des étudiants.
Traditionnellement, la courbe d'apprentissage est considérée comme étant non linéaire et conforme à la « loi des rendements décroissants », c'est-à-dire que le temps passé pour étudier procure un gain marginal de plus en plus faible en termes de réussite.

Dans la relation étudiant-enseignant, chaque intervenant joue un rôle actif. L'étudiant a un comportement dual (Bishop, 2006). Il possède un éventail de choix entre fournir un effort maximal pour l'apprentissage et l'abandon complet de la formation. Cette décision est influencée par sa préférence pour les options extérieures : le loisir, le travail rémunéré, etc. L'étudiant est alors considéré comme un consommateur/investisseur qui choisit d'allouer un temps optimal pour l'apprentissage. Mais il est également considéré comme un producteur à part entière de l'apprentissage, qui reçoit les instructions, les conseils, et les évaluations de ses enseignants. On peut alors noter la fonction d'utilité qu'il reçoit de l'apprentissage

$$U_{étudiant} = U(w_{étudiant}, t_{étudiant})$$

où $t_{étudiant}$ correspond à l'effort d'apprentissage qui dépend du temps alloué à d'autres activités et $w_{étudiant}$ représente une fonction de récompense de l'apprentissage. L'effort est coûteux pour l'étudiant. Il dépend du type, du rythme, de la qualité, de la durée des activités d'apprentissage (Leibenstein, 1980).
La fonction de gain correspond à un gain non pécuniaire, c'est-à-dire un certain de degré de « joie » de l'apprentissage. Elle mesure le niveau de résultat perçu lié au fait d'apprendre et de montrer le niveau d'apprentissage obtenu après un examen, par exemple la position dans un classement par les notes. Ce paramètre reflète l'idée que l'étudiant n'a pas une directe perception de sa réussite, mais seulement à travers le filtre indirect des notes. La notation peut être normale, sévère, ou clémente.

Il est important de relever qu'à ce stade de la modélisation proposée par Correa et Gruver (1987), ces hypothèses impliquent que le comportement de l'enseignant influence l'utilité de l'étudiant uniquement à travers la réussite perçue, c'est-à-dire la note. Par conséquent, les autres caractéristiques de l'enseignant, comme la possibilité de rendre un enseignement plus attrayant ou l'interaction des modalités de l'évaluation en tant que dispositif de vérification des connaissances acquises avec le comportement des étudiants, ne sont pas considérées dans cette partie du modèle. Nous y reviendrons plus tard dans notre exposé, lorsque nous évoquerons la prise en compte des ajustements comportementaux à l'équilibre et du contrôle des connaissances.

Enfin, en ce qui concerne le comportement de l'enseignant, ce dernier joue également un rôle dual. Il est un consommateur/investisseur et sa décision d'enseignement dépend de ses options extérieures. Un enseignant du supérieur répartit son temps entre la recherche, les tâches administratives et l'enseignement. Dans ce modèle, l'instruction associe dans la même dimension le contrôle et la délivrance des connaissances. L'enseignant est également un producteur à part entière de l'apprentissage, ce qui lui permet d'avoir une récompense de l'effort fourni.

À l'instar de la fonction d'utilité de l'étudiant, on peut noter

$$U_{enseignant} = U(w_{enseignant}, t_{enseignant})$$

où chaque paramètre correspond à ce qui a été décrit pour l'étudiant.

Dans cette relation, chaque individu choisit le niveau d'effort optimal pour l'apprentissage qui maximise son utilité étant donné la contrainte d'effort pour l'apprentissage pour acquérir (délivrer) des connaissances pour l'étudiant (enseignant), le niveau de récompense perçu par l'apprentissage, et la courbe d'apprentissage. Cela correspond à un programme d'optimisation de l'utilité sous une contrainte technologique.

Pour l'étudiant, l'effort optimal d'apprentissage dépend du niveau d'instruction réalisé par l'enseignant, du niveau de récompense et des préférences pour la réussite et pour l'apprentissage.
La réaction de l'étudiant à un niveau d'instruction plus élevé, c'est-à-dire à des activités plus fréquentes, à des demandes d'acquisition de connaissances plus importantes, à des examens plus difficiles, se traduit par une élévation de l'effort si et seulement si la préférence pour les activités extérieures à l'apprentissage est inférieure au degré de substitution de la courbe d'apprentissage entre l'effort de l'étudiant et l'effort de l'enseignant.
En ce qui concerne la réaction à une notation plus sévère, qui se traduirait par un niveau de récompense perçu comme étant moins important, tout dépend du degré de préférence pour l'apprentissage. Ce qui peut être incitatif pour un étudiant peut être contreproductif pour un autre. Les mêmes conclusions s'appliquent pour l'enseignant.
Dans une certaine mesure, la forme de la fonction de récompense au travers des notes induirait également une augmentation de l'effort de l'apprentissage (Bishop, 2014). Cette intuition a permis le développement de travaux sur la forme des notes et la réaction des individus à celle-ci (Costrell, 1994 ; Bonesrønning, 2004). Cependant, cette littérature ne permet pas d'analyser le processus de l'évaluation et par conséquent elle ne suffit pas à illustrer le propos de ce chapitre.

En définitive, Brown et Sacks (1981) ainsi que Correa et Gruver (1986) notent que les préférences individuelles semblent être le principal déterminant de l'effort et du niveau de

l'apprentissage, ce que nous pourrions interpréter comme étant le goût pour l'apprentissage. L'appétence pour les études pour les étudiants ou le goût pour la transmission de connaissances pour l'enseignant seraient les uniques déterminants de la réussite.

Ce résultat interpelle quant à la faible capacité à modifier l'apprentissage autrement qu'en modifiant les préférences. Une manière de contourner cette prédiction est de prendre en compte les interactions qui se réalisent au sein de la relation. En effet, les niveaux d'effort d'apprentissage et d'instruction de l'étudiant et de l'enseignant se modifient en réaction à ce que décide l'autre individu.

Dans le précédent développement, le niveau d'effort de l'étudiant est réalisé en tenant compte du niveau d'effort anticipé de l'enseignant et réciproquement. Or, il n'y a pas d'éléments supplémentaires qui permettent de penser que le niveau estimé d'effort de l'autre agent est celui qui se réalise véritablement. Cette situation conduit donc à des réajustements dans les comportements des deux parties.

Ces réajustements caractérisent un comportement interactif sans coopération. De surcroît, il est possible de considérer l'émergence d'un comportement coopératif, dans le cas où l'étudiant et l'enseignant n'observent pas les valeurs prises par l'effort. Ainsi, chaque partie tente d'influencer directement l'autre dans la détermination de l'effort et négocie un accord conjoint.

b. Lorsque l'éducation est un bien public, la coopération permet d'atteindre un optimum social

D'un point de vue théorique, Correa et Gruver (1986) ont démontré l'existence d'un équilibre en situation d'interactions dans la relation d'instruction. Le processus est le suivant : l'étudiant va choisir un niveau d'effort d'apprentissage étant donné un certain niveau d'instruction, la difficulté des activités, de la discipline demandée par l'enseignant, etc. ; s'il anticipe un niveau d'instruction différent, à la hausse ou à la baisse, cela se traduit par une variation équivalente de la courbe d'apprentissage, puisque celle-ci dépend des caractéristiques des deux parties.

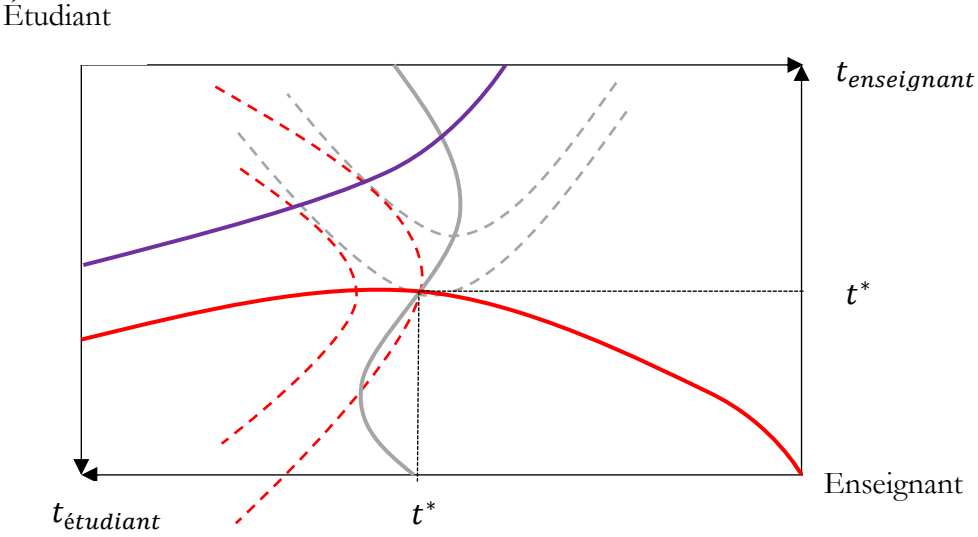
Tous ces ajustements successifs se matérialisent dans une fonction de réaction qui contient l'ensemble des niveaux d'efforts optimaux d'apprentissage qui dépendent des niveaux d'instruction de l'enseignant. Le même processus se manifeste pour l'enseignant. L'équilibre est obtenu à l'intersection des deux fonctions de réactions.

La stabilité de cet équilibre est étudiée grâce à un processus d'ajustement à la Cournot, c'est-à-dire dans lequel l'étudiant et l'enseignant modifient respectivement leur niveau d'apprentissage et d'instruction en présumant le fait que l'autre partie maintient sa stratégie. Les auteurs démontrent que l'équilibre est stable si et seulement si la valeur accordée à l'effort d'apprentissage par

l'étudiant, i.e. sa préférence, est « proche » de la valeur accordée à la réussite scolaire, et que la valeur accordée à l'effort d'instruction par l'enseignant est « proche » de la valeur accordée à la réussite scolaire.

En définitive, une condition nécessaire de la stabilité de l'équilibre requiert que la valeur accordée à l'apprentissage et à l'instruction soit similaire en magnitude.

Graphique 1 : L'équilibre coopératif de la relation enseignant-étudiant



De nouveau, c'est le goût de l'apprentissage qui détermine les réactions comportementales. L'enjeu pour l'analyste est alors de déterminer le niveau de préférence pour inciter à choisir le meilleur « couple » d'effort d'apprentissage et d'instruction pour parvenir à un niveau de réussite prédéterminé.

Cependant, les auteurs ne s'avancent pas quant au niveau de réussite scolaire qui serait socialement acceptable. Pour le déterminer, il faudrait inclure un critère de jugement sur ce qu'il est souhaitable d'atteindre.

Néanmoins, Correa et Gruver (1986) ont achevé leurs démonstrations en incluant la dimension de coopération et de négociation, afin de comparer si cette situation était préférable à une situation d'interaction sans coopération. Il est possible d'analyser la situation grâce à une représentation graphique (graphique 1).

Les courbes vertes et rouges en trait plein correspondent aux fonctions de réaction de l'enseignant et de l'étudiant respectivement. Leur intersection forme l'équilibre d'interaction sans coopération ($t^*_{étudiant}, t^*_{enseignant}$).

À ce niveau d'effort individuel correspondent des niveaux de satisfaction représentés par les courbes d'indifférences en trait pointillé. Il existe donc des situations qui procurent un niveau de

satisfaction plus élevé pour les deux parties : ce sont toutes les situations qui se situent dans l'ellipse formée par l'intersection des deux courbes d'indifférence qui passent par l'équilibre non coopératif. En plus de procurer un niveau de satisfaction plus important, ces situations procurent également un niveau de réussite plus important. Mais ce ne sont pas des « allocations » d'efforts qui permettent de former un équilibre stable, puisque la stabilité se situe au niveau de l'interaction sans coopération.
En revanche, dans ce cadre, il existe un ensemble de couples d'efforts représenté par la courbe violette qui désigne la courbe des contrats, c'est-à-dire l'ensemble des allocations Pareto-optimales. Ces allocations s'obtiennent dès lors qu'il y a une concertation entre les deux individus. Graphiquement et mathématiquement, la situation sans coopération, est une situation inférieure à celle qui prévaudrait dans le cas parétien. L'intersection entre l'ellipse et la courbe des contrats contient alors toutes les solutions d'équilibre relatives à la négociation pour l'étudiant et l'enseignant sur ce que serait un niveau socialement acceptable d'effort d'apprentissage et d'effort d'instruction pour obtenir un résultat scolaire plus important.

C'est un cas spécifique obtenu dans la littérature de l'allocation des ressources pour la production d'un bien public. Il est généralement admis que des optimisations indépendantes conduisent à des résultats qui sont inférieurs à l'optimum de Pareto. La raison pour laquelle l'interaction sans coopération conduit à des résultats scolaires inférieurs est que chaque individu considère les bénéfices pour eux-mêmes et pas les externalités qui bénéficient aux autres.
Selon les auteurs, une stratégie de menace serait susceptible d'opérer un déplacement suffisamment incitatif vers une solution supérieure à une situation sans coopération. Dans ce cas, il est légitime de penser que des interactions entre l'étudiant et l'enseignant qui permettraient de dialoguer sur les attentes réciproques sur le niveau d'effort et sur les dispositifs de contrôle de connaissances seraient favorables à l'atteinte de résultats scolaires plus importants qu'en l'absence de négociation.

Malgré la richesse de ce premier modèle, des zones d'ombre demeurent sur la répartition des activités de l'enseignant, notamment l'arbitrage entre son temps passé à transmettre des connaissances et à évaluer le niveau des étudiants. Néanmoins, ce modèle permet d'aboutir à une dimension importante de la relation enseignant-étudiant : la coopération sur le niveau d'effort d'apprentissage et sur le niveau d'instruction peut conduire à une situation meilleure que celle sans négociation. Sans davantage de précision sur les termes de la négociation, il apparaît possible d'améliorer la réussite des étudiants en faisant dialoguer les deux parties de la relation.

c. L'existence de divergence d'intérêt sur l'éducation conduit à l'existence d'une relation d'agence

Il existe dans la littérature plusieurs auteurs qui tentent de préciser les termes de la relation étudiant-enseignant. Comme nous l'avons décrit précédemment, il existe des interdépendances, des ajustements individuels, entre les deux parties, qu'il s'agit d'affiner pour ouvrir la « boîte noire » de cette relation. Ces modèles partent de l'hypothèse que l'éducation n'est plus un bien public, mais qu'il procure des bénéfices individuels. Par conséquent, il existe des divergences d'intérêt sur ce que procure ce bien.

Pour aligner les intérêts des deux parties, il faut résoudre un problème d'incitation. Dans ce cadre, les auteurs ont davantage conceptualisé le dispositif d'évaluation des connaissances, qui a un double rôle : mesurer le niveau de connaissance et inciter à l'effort d'apprentissage.

Dans un modèle principal-agent, Zubrickas (2015) part du principe qu'il existe une vision dichotomique de l'éducation : d'une part elle assure la formation du capital humain, d'autre part elle signale les compétences acquises sur le marché du travail. Les enseignants sont concernés par la première vision, tandis que les étudiants ne s'intéressent qu'au signal. Dans ce contexte, les étudiants traitent la connaissance comme une chose abstraite, éphémère et coûteuse, comparativement aux opportunités extérieures. C'est ce type d'attitude qui crée un conflit d'intérêts entre l'enseignant et l'étudiant.

Dans cette relation, l'enseignant fait face à un problème de sélection des étudiants qui auraient des intérêts alignés avec les siens. L'enseignant prodigue un enseignement à une classe d'étudiants avec des aptitudes différentes pour un sujet donné. Son objectif est de délivrer des connaissances. L'aptitude est une information personnelle et plus l'aptitude est importante, plus l'effort d'apprentissage est faible pour atteindre le niveau de connaissances.

L'enseignant possède alors d'un dispositif, l'examen, qui lui permet d'évaluer le niveau de connaissances atteint par les étudiants. Par extension, l'enseignant définit une règle de notation qui assigne des scores au résultat de l'examen, c'est-à-dire au niveau de connaissances. Par hypothèse, l'examen est « parfait », c'est-à-dire qu'il y a une correspondance parfaite sur le niveau de connaissance évalué et la note donnée. La conséquence de ce postulat est que l'évaluation de la performance est uniquement assurée par les notes. Cette hypothèse permet de simplifier la modélisation, cependant elle empêche la formulation d'une description plus précise du contrôle des connaissances.

L'aptitude est distribuée dans la population sur un espace $[\theta_L, \theta_H]$, l'indice L mesurant le niveau d'aptitude faible et l'indice H le niveau d'aptitude élevé. Atteindre un résultat d'examen x, c'est-à-dire avoir un certain niveau de connaissance (dans le modèle canonique, cela pourrait

correspondre à la courbe d'apprentissage), coûte à l'étudiant un effort $\frac{C(x)}{\theta}$. L'enseignant utilise le barème de note r pour inciter à l'effort.

La relation se développe de la façon suivante : l'enseignant désigne d'abord un barème, qui fournit une relation d'allocation entre le résultat à l'examen et la notation, $\{x(\theta), r(\theta)\}$; puis l'étudiant montre son aptitude, ou de manière équivalente, choisit le contrat entre l'examen et un certain barème. L'enseignant maximise son utilité espérée $V(x(\theta)) = x$ qui est linéaire sur le résultat d'examen. L'étudiant maximise son utilité espérée $\phi(r) - \frac{C(x)}{\theta}$, qui ne dépend que des notes. Lorsque le modèle est résolu, la règle de notation optimale, qui est considérée ici comme le contrat optimal d'incitation, est celle qui maximise le niveau de connaissances d'un étudiant dont le but est d'avoir la meilleure note qui signale ses compétences sur le marché du travail.

Les conclusions de ce modèle vont bien au-delà du champ d'investigation considéré dans ce chapitre : en effet, ce modèle est utilisé pour prédire le phénomène d'inflation des notes, c'est-à-dire la situation où la quantité de notes est excessive pour juger la qualité des candidats sur le marché du travail. Ce modèle prédit également une compression des notes, c'est-à-dire une forte homogénéité des notes données par les enseignants. Il fournit des explications sur la réduction du temps passé dans l'enseignement par les étudiants. Enfin, plus un enseignant est pessimiste sur le niveau d'aptitudes des étudiants, plus il sera indulgent dans sa notation.

Néanmoins, il semble pertinent pour qualifier le risque d'aléa moral sur les attentes des étudiants sur l'enseignement, au même titre qu'il existe un aléa moral sur les attentes des enseignants sur l'effort des étudiants. Le principal apport de ce type d'analyse est que ces divergences d'intérêt conduisent à une mésentente sur les résultats de l'éducation. Considérant le faible taux de réussite à l'université, il est permis de penser que les attentes de l'enseignement supérieur ne coïncident pas avec celles des étudiants. De surcroît, le contrat qui est conclu entre les deux parties amènerait à des résultats inefficaces. Nous faisons le postulat qu'un des paramètres de ce contrat, à savoir le contrôle des connaissances, n'est pas efficacement calibré pour permettre d'atteindre un équilibre socialement partagé.

Tandis que pour le modèle canonique, il n'existait pas de précision sur les modalités de l'évaluation, il est possible d'apprécier les effets de l'examen sur la relation enseignant-étudiant. Pour cela, il faut se situer au niveau de la théorie économique des standards d'éducation (Becker et Rosen, 1992 ; Michaelis et Schwanebeck, 2015). À notre connaissance, c'est l'approche qui a proposé de modéliser le dispositif d'évaluation des acquis et ses interactions avec le comportement d'apprentissage des étudiants.

Dans ces modèles, plutôt appliqués à un contexte d'examen national, comme le baccalauréat, la relation d'agence n'est pas explicitement évoquée, mais elle est sous-jacente. Un décideur public

cherche un examen qui maximise une fonction de bien-être social, sous-entendu le niveau de connaissance (le capital humain), en tenant compte de l'aversion à l'effort d'apprentissage des étudiants. L'analyse se concentre sur l'arbitrage entre l'effet d'incitation et de classement qu'un changement dans les modalités d'examen peut produire sur le comportement des étudiants.

Le cadre d'analyse commence avec le résultat à l'examen, q, qui est la somme de deux composants, les caractéristiques de l'étudiant et les caractéristiques de l'examen :

$$q = \mu + \epsilon$$

L'étudiant modifie la variable μ en apprenant et en s'exerçant, ainsi qu'en utilisant ses aptitudes personnelles pour maîtriser les connaissances qu'on attend de lui. Comme pour le modèle canonique, l'étudiant décide du niveau d'effort dédié à l'apprentissage.

En revanche, il n'a pas de possibilités pour influencer le terme ϵ, qui représente un aléa sur les évènements qui se produisent durant l'examen. Ces aléas relèvent d'une part, de la qualité de l'évaluation, par exemple la capacité de mesurer convenablement le niveau d'apprentissage, ou sa forme, par exemple le fait de répondre à un questionnaire à choix multiple, d'autre part de l'état physique et émotionnel de l'étudiant, son niveau de stress ou d'anxiété.

Par conséquent pour un test donné à un étudiant quelconque, on note

$$\epsilon_{i_t} = v_t + u_{i_t}$$

Le résultat issu de l'aléa n'est pas connu à l'avance, mais les étudiants connaissent sa distribution $F(\epsilon)$ et sa densité $f(\epsilon)$ quand ils décident du niveau d'effort destiné à apprendre.

Dans ce contexte, il existe une sorte de courbe d'apprentissage qui est représentée par la fonction d'effort $c(\mu)$. Cependant, elle n'a pas les mêmes propriétés que dans le cadre canonique, mais deux postulats sont conservés. L'effort procure de la désutilité, puisque le temps passé à apprendre n'est pas dédié à d'autres activités (loisirs ou travail rémunéré). Par conséquent, le coût marginal de l'effort est positif. De plus, le rendement de l'apprentissage est décroissant, plus on réalise un effort pour apprendre, plus le gain en termes de connaissances est réduit, ce qui se traduit par un coût d'effort de plus en plus élevé. Le coût marginal de l'effort est donc croissant.

Le résultat de l'examen conduit à deux situations : réussir ou ne pas réussir. C'est à ce niveau qu'intervient le barème de « récompenses », c'est-à-dire les notes associées à un niveau de réussite. On note R_1, la note associée à la réussite, et R_2 la note associée à l'échec.

Comme ces modèles s'adressent davantage à l'économie des examens à un niveau national, le décideur public fixe un seuil de connaissance pour lequel on certifie la réussite à une limite $\bar{q}$. Il existe donc une probabilité qu'un étudiant réussisse ou échoue à l'examen, notée :

$$p_1 = p(q \geq \bar{q})$$
$$p_1 = p(\epsilon \geq \bar{q} - \mu)$$
$$p_1 = 1 - F(\bar{q} - \mu)$$

L'objectif de l'étudiant est de maximiser son utilité espérée de réussir sous contrainte de coût d'effort,

$$EU = p_1.R_1 + (1 - p_1).R_2 - c(\mu)$$
$$EU = R_2 + p_1.(R_1 - R_2) - c(\mu)$$
$$EU = R_1 - F(\bar{q} - \mu).(R_1 - R_2) - c(\mu)$$

En choisissant l'effort qui maximise son utilité, l'étudiant compare le temps nécessaire à atteindre un niveau de connaissance qui accroît ses chances de réussir à dépasser le niveau requis avec le différentiel de récompense associé à la réussite.

Dans ce modèle initialement proposé par Becker et Rosen (1992), le changement du niveau de validation des connaissances, $\bar{q}$, conduit à des résultats ambigus. Si l'on pense que les standards sont bas, les accroître conduit les étudiants à travailler plus, mais seulement jusqu'à un certain point. Passé ce niveau critique, la situation s'inverse et augmenter le niveau d'exigence peut se transformer en une réduction de l'effort. Par conséquent, mettre des objectifs trop bas ou trop élevés conduit les étudiants à ne pas fournir l'effort adéquat. Les auteurs notent que pour une population étudiante hétérogène, mettre un niveau d'acquisition de connaissances à un niveau national peut être questionnable : quel serait le niveau adéquat pour inciter tous les étudiants à fournir les efforts que l'on attend d'eux ?

Hormis, Michaelis et Schwanebeck (2016) qui s'intéressent à la réaction comportementale des étudiants à différentes variantes d'un examen de rattrapage (avec droit d'inscription, niveau requis au rattrapage différent de la première session, temps passé entre chaque examen, point malus), il ne semble pas exister d'autres tentatives pour examiner les méthodes d'évaluation à l'intérieur des classes, qui sont des manières plus régulières d'examiner l'apprentissage des étudiants. Par conséquent, c'est un champ de recherche qui demeure ouvert pour de futures investigations.

Ce qu'il faut retenir de cette approche est que l'évaluation conditionne l'effort d'apprentissage des étudiants, mais elle peut également causer des effets négatifs. Il s'agit alors de trouver un critère qui permettrait de désigner ce qu'il est souhaitable d'atteindre socialement.

d. La possibilité de négocier le contrat optimal dans la relation étudiant-enseignant

Nous avons tenu compte de toutes ses considérations pour inscrire l'idée que la relation étudiant-enseignant est interconnectée et que le changement dans les règles d'examen peut provoquer des réactions comportementales dans les deux parties.
Nous terminons cette revue de la littérature théorique, en utilisant les apports de la négociation dans le cadre de la théorie des contrats. De cette manière, nous faisons le parallèle avec le modèle canonique, qui explicitait la relation de coopération, mais pour la production d'un bien public. Il est judicieux de déterminer si cette conclusion s'applique en présence d'aléa moral.

Une hypothèse traditionnelle de la relation d'agence est que le principal offre un contrat « à prendre ou à laisser », c'est-à-dire que le principal possède tout le pouvoir de négociation. Pourtant, dans des situations concrètes, les deux parties ont des marges de manœuvre pour négocier les termes du contrat. Pour Pitchford (1998), ces situations recouvrent les cas où la responsabilité de l'agent est limitée, par exemple pour le remboursement des emprunts quand la capacité de paiement est faible. Dans ce cas, le pouvoir de négociation est représenté par la variation de l'utilité de réserve de l'agent, c'est-à-dire en faisant varier ses opportunités extérieures. Balckenborg (2001) utilise une approche différente en recourant à la solution de négociation à la Nash (Nash, 1950). Demougin et Helm (2004) étudient un jeu d'offre alternative avec aléa moral à la Rubinstein (1982). Enfin, Dittrich et Städter (2015) emploient la solution de négociation à la Kalaï-Smorodinsky (Kalaï et Smorodinsky, 1975).

La littérature sur la négociation dans un cadre d'aléa moral suggère qu'il existe une courbe des contrats qui procure une frontière Pareto-optimale des utilités du principal et de l'agent, à chaque fois qu'intervient une distribution du pouvoir de négociation entre les deux parties.
En revanche, le choix d'une solution théorique de négociation ne doit pas être une décision arbitraire puisque la préférence pour l'une ou l'autre solution peut sensiblement modifier les résultats de la négociation (Demougin et Helm, 2004) : dans le cas du changement de l'utilité de réserve et dans le jeu d'offre alternative, il n'y a pas d'effet sur l'éventail de choix du contrat optimal, i.e. la courbe des contrats Pareto-optimale n'est pas modifiée ; dans le cas de l'utilisation des solutions de négociation à la Nash ou à la Kalaï-Smorodinsky, la courbe des contrats peut être modifiée.

Ce que nous retirons comme principal apport de cette littérature est qu'il est possible de donner un pouvoir de négociation à l'agent, en présence d'aléa moral, qui permet d'atteindre un optimum social plus efficace qu'en situation sans négociation.

Ce résultat est donc similaire à ce qui a été déterminé dans le cadre canonique de l'économie de l'éducation. Il est donc raisonnable de penser qu'il existe des marges de manœuvre pour configurer une relation d'apprentissage où l'étudiant entrerait en négociation avec l'enseignant. Les termes de la négociation peuvent porter sur la règle d'examen, sur les notes attribuées lorsque l'examen est validé ou pas, sur le niveau d'effort attendu de l'étudiant pour acquérir des connaissances.

Il nous est alors permis de penser que dans le cadre d'une relation étudiant-enseignant, la négociation de la règle d'examen pour quantifier le niveau de connaissance obtenu serait une façon d'améliorer les positions des deux individus, et de contribuer à l'élévation du niveau de réussite des étudiants, y compris dans le cas où les deux parties ont des divergences d'intérêt sur l'éducation.

2. Des pistes ouvertes par la littérature appliquée pour évaluer la négociation des règles d'évaluation des connaissances

a. En économie, une recherche focalisée sur l'effet de la notation

En économie, la littérature appliquée qui rentre dans le champ de la relation enseignant étudiant est moins importante que celle qui concerne l'investissement dans le capital humain. Elle s'est d'abord appliquée à étudier les facteurs de production de l'apprentissage qui ont été présentés dans la première partie de la présentation des modèles théoriques, en estimant leur effet sur la production de la réussite éducative (Rivkin et al., 2005). Ces travaux confirment les principales intuitions des modèles : les principaux déterminants de la réussite sont les préférences individuelles.

L'approfondissement de cette approche a été tourné vers l'influence des notes sur le comportement d'apprentissage (Bonesrønning, 2004). Dans ce cadre, les notes sont perçues comme des formes d'incitations à l'effort d'apprentissage. Il s'agit alors de déterminer la forme optimale de la notation et du barème.

En revanche, il ne semble pas exister de travaux empiriques qui se focalisent explicitement sur les pratiques d'examen des connaissances et leurs effets sur l'effort d'apprentissage. De plus, sauf erreur de diagnostic de notre part, il ne semble pas qu'il existe actuellement de recherche appliquée portant sur la négociation de la règle de vérification de l'effort dans un cadre de relation d'agence.

b. La prise en compte de la dimension partenariale dans la recherche en sciences de l'éducation

Dans la littérature, ce sont des chercheurs en sciences de l'éducation qui ont pris l'initiative de déterminer dans quelle mesure le partenariat dans l'évaluation permettait de dépasser les faibles résultats académiques des étudiants. Les méthodes d'investigation employées sont à mi-chemin entre l'étude de cas et la recherche-action en recourant à des dispositifs d'analyse qualitatifs. Ces contributions tentent de développer un nouveau paradigme qui semble répondre aux préoccupations de l'enseignement supérieur et elles remettent en question les rapports de pouvoir autour du contrôle des connaissances.

L'ensemble de ces travaux partent du constat d'un manque d'efficacité dans le modèle traditionnel de l'évaluation dans l'enseignement supérieur. Il est décrit comme autoritaire et implique la possession inégalitaire de l'exercice du pouvoir : le corps enseignant décide ce que les étudiants doivent apprendre ; il détermine le critère d'évaluation ; il évalue. Les étudiants sont évincés de chaque étape.
Cela provoquerait des conséquences négatives : un sentiment d'injustice ; un manque de fiabilité dans ses propres compétences ; une stratégie qui conduit à vouloir obtenir une bonne note et sans passer par l'apprentissage (Falchikov, 1986). Les étudiants ne semblent pas suffisamment être des parties prenantes de leur cursus et ne prennent pas suffisamment de responsabilités dans l'apprentissage (Stefani, 1996).
Il faut donc les équiper pour qu'ils soient flexibles, adaptables et préparés à être responsables de leur apprentissage tout au long de la vie, y compris sur leur futur lieu de travail. L'évaluation permet de jouer un rôle dans leur estime de soi à partir du moment où il est possible de situer soi-même les critères de réussite, de se poser soi-même les questions et d'y répondre en autonomie.

Dans ce contexte, l'idée s'est développée de rendre plus démocratique l'environnement des cursus de formation en faisant intervenir plus activement les étudiants à divers degrés, notamment sur les paramètres de l'évaluation : quelle forme de l'évaluation ? Comment évaluer ? Cette démocratie d'enseignement permet de faire se rencontrer les deux parties de la relation et de créer un dialogue pour lever le voile sur les critères de jugement à l'œuvre à l'université, c'est-à-dire d'effacer l'implicite des attendus. Cet environnement incite les étudiants à verbaliser leur ressenti et leurs attentes de la formation.

Le concept clé est celui de partenariat, conçu comme « un processus collaboratif et de réciprocité dans lequel tous les participants ont l'opportunité de contribuer équitablement, mais dans des conditions différentes, au programme ou à la conception pédagogique, à la prise de décision, à

l'investigation et à l'analyse » (Bovill, et al., 2011). Dans ce contexte, il existe des opportunités pour donner du pouvoir de négociation aux étudiants pour intervenir dans l'évaluation de manière collaborative avec l'équipe pédagogique.

Cette pratique du partenariat permet de changer les rôles : l'enseignant échange sa position autoritaire de surveillant avec celle d'accompagnant dans l'apprentissage ; l'étudiant troque sa fonction de consommateur passif du savoir à celle de contributeur actif à l'apprentissage.

C'est une collaboration dans le sens où les deux parties travaillent pour un objectif partagé : développer l'apprentissage par la compréhension et la création des objectifs et des critères d'évaluation. L'idée est de renforcer des compétences de contrôle et de jugement tout en développant une compréhension et un intérêt pour le processus d'examen des connaissances. Par conséquent, cela augmente la littéracie de l'évaluation et les étudiants sont capables de s'approprier l'évaluation (Rust et al., 2003). *In fine*, l'objectif de ce dispositif est d'améliorer la réussite scolaire parce que les étudiants sont davantage réceptifs aux attendus de l'évaluation dans une formation d'enseignement supérieur.

Dans ce contexte, m'étudiant n'est pas nécessairement responsable de l'évaluation, mais il collabore dans le processus et détermine avec l'enseignant ce qui sera évalué et par quel outil. Dans ce partenariat, l'évaluation est alors coconstruite, c'est-à-dire que l'évaluation engendre « un processus actif en évolution utilisé par l'apprenti comme la partie d'un tout créatif où la compréhension du dispositif n'est plus un produit à accumuler » (Deeley, 2014). De plus, selon Dochy et al. (1999) le partenariat peut aboutir à une forme de co-évaluation : la participation des étudiants et de l'équipe enseignante dans le processus d'évaluation est un moyen de donner l'opportunité de s'auto évaluer tandis que le tuteur garde un contrôle sur l'évaluation finale.

La culture du partenariat peut s'effectuer de plusieurs manières, mais il faut donner du crédit au partenariat, c'est-à-dire le récompenser. Cela se traduit concrètement par la notation consécutive à la construction des règles.

Au niveau conceptuel, les contributions sont raisonnablement positives sur les effets que le partenariat peut prodiguer sur la pédagogie et sur les capacités d'apprentissage. Cependant, il y a de nombreuses zones d'ombres sur le niveau opérationnel. Nous faisons ci-après la recension des principales publications de la recherche sur le partenariat dans l'évaluation.

c. Les effets de négociation de la règle de l'évaluation

Falchikov (1986) a expérimenté l'évaluation partenariale avec 48 étudiants répartis dans trois groupes de première année de licence de biologie à partir d'un processus itératif sur les règles et la

grille d'évaluation d'un partiel d'introduction à la psychologie, en Écosse. Les chargés de travaux dirigés ont d'abord mis en commun leurs règles, puis ils ont demandé aux étudiants de proposer leurs propres critères de notation, par un processus de consensus. La grille d'évaluation finale contenait à la fois les critères des étudiants et ceux des enseignants. La grille était commune entre les groupes et était utilisée pour de l'auto-évaluation des étudiants et l'évaluation par les tuteurs. La note finale était ensuite négociée individuellement.

La question de recherche était de déterminer si la création d'une grille commune permettait de situer au même niveau le résultat des étudiants qu'il soit évalué personnellement ou par un évaluateur externe, c'est-à-dire de savoir si la responsabilité de l'évaluation donnée aux étudiants engendrait des comportements de surnotation par rapport au cas où l'enseignant notait seul.

Sur le processus de cocréation, les étudiants étaient plutôt satisfaits : ils ont perçu une meilleure équité de traitement que dans un contexte sans collaboration. En revanche, ils ont été insatisfaits de la responsabilité de notation, parce que la grille était difficile à utiliser, et qu'elle n'était pas suffisamment pertinente pour juger le travail réalisé.

Sur les résultats de la notation, les étudiants se sont jugés plus sévèrement que les enseignants dans la majorité des cas (en moyenne l'écart avec l'enseignant est de 7,67 points sur 100).

L'auteure semble noter que malgré la difficulté de la prise en charge de l'évaluation par les étudiants, ils ont réalisé le travail. Elle constate une progression des étudiants, car cela les a rendus davantage critiques sur ce qu'il fallait apprendre et sur la manière de rédiger les réponses au partiel. Ce processus a également permis de développer leur capacité de jugement sur leurs propres résultats. Ils ont davantage eu de réflexions sur les enjeux de l'apprentissage. Par conséquent, l'évaluation collaborative a renforcé la responsabilité et l'autonomie des étudiants.

Stefani (1998) a mobilisé le partenariat évaluatif dans une classe de trente étudiants de première année de master, répartis dans six groupes de cinq d'un module de formation à l'ingénierie environnementale. La tâche était de préparer un dossier de recherche documentaire thématisé. Ce genre d'activité était traditionnellement auto-évalué et évalué par les enseignants. En revanche, l'équipe pédagogique notait que les précédentes promotions n'arrivaient pas à bien réaliser ce travail : « les étudiants ne sont pas tous conscients de ce qu'on attend d'eux et quels sont les critères d'examen ».

L'intervention était la suivante : chaque groupe définissait la programmation du dossier, les ressources à mobiliser, la répartition des tâches, les compétences de chaque étudiant et les difficultés prévues. Le processus de notation était négocié entre l'enseignant et les étudiants : quels critères d'excellence retenir ? Quel barème de notation ? Chaque groupe tenait un journal de bord qui a été ensuite analysé par la chercheuse.

Les conclusions de ce travail sont que les étudiants ont trouvé la tâche plus difficile que si les objectifs étaient définis pour eux, mais ils développent un sens plus accru des responsabilités parce qu'ils ont atteint ou dépassé leurs propres objectifs. L'auteure conclut que les résultats ont été plus satisfaisants que les précédentes promotions et que les dossiers étaient de meilleure qualité qu'auparavant (sans préciser l'ampleur du résultat).

Garside et al. (2009) ont reformulé le dispositif d'évaluation d'un module de formation pour les étudiants en deuxième année d'école d'infirmier. 157 étudiant.e.s ont participé à ce processus. L'évaluation consistait à choisir trois modes d'examen des connaissances, entre une dissertation de 2000 mots, un oral de 15 minutes, ou un devoir surveillé de 90 minutes comptant pour la moitié de la note finale (un partiel « *Objective Structured Clinical Examination [OSCE]* »[19] comptait pour l'autre moitié). Un questionnaire auto administré a été développé pour collecter le ressenti des étudiants. Le mode d'évaluation par la dissertation a été choisi par 90 % des étudiant.e.s, parce que c'était « plus simple pour programmer l'apprentissage » et « plus confortable », tandis que l'examen oral et le devoir surveillé étaient davantage perçus comme stressant.
Donner le choix aux étudiants a été ressenti très positivement par la quasi-intégralité des étudiants. Ces derniers ont ressenti être en meilleure condition pour réussir l'examen : le fait d'avoir choisi la manière d'être évalué, en choisissant ce qui était le plus « agréable » a permis d'être en confiance quant aux connaissances acquises et dans leurs performances pour réussir la tâche qu'ils s'étaient eux-mêmes assignée. Cependant, quantitativement, les auteurs n'ont pas mesuré le gain en termes de note par rapport à une situation sans choix d'examen.

Deeley (2014) a expérimenté la notation conjointe à travers la négociation de la grille d'évaluation d'un cours optionnel de politique publique de troisième année et licence et de première année de master d'une université écossaise. La négociation est intervenue sur la notation d'un examen oral à la fin du semestre portant à la fois sur des connaissances issues du cours magistral et sur l'expérience vécue au cours d'un stage dans un centre de travail social. La note de ce module comptait pour 10 % de l'unité d'enseignement et 2,5 % du diplôme. Les étudiants s'autoévaluaient et l'enseignant donnait une note sur la prestation orale. À la suite de l'examen, une discussion s'engageait pour négocier la note. Si la négociation ne débouchait pas, l'enseignant avait le dernier mot. Des entretiens individuels ont été menés pour évaluer qualitativement les effets de cette structuration de l'évaluation.
Les étudiants ont trouvé que c'était une aide majeure dans leur apprentissage parce que cela a permis de confirmer leur appréciation individuelle du niveau de connaissances acquis. La

[19] C'est un examen conçu pour évaluer les compétences cliniques comme la communication, le diagnostic médical, la prescription, les techniques de manipulation, le positionnement radiographique, et l'interprétation des résultats d'examens médicaux.

négociation de la note a permis de développer des compétences de communication et a renforcé leur motivation dans l'apprentissage, afin de montrer ce qu'ils étaient capables de réaliser et de développer des arguments pour convaincre l'enseignant. L'évaluation négociée a été perçue comme plus équitable que la simple auto-évaluation. La stratégie de vouloir s'attribuer des notes très hautes a été rejetée collectivement, parce qu'ils ont pensé que cela irait à l'encontre de leur apprentissage, que cela contreviendrait à la nature de l'intervention et serait irrespectueux envers l'enseignant.

Deeley et Brown (2014) ont rédigé un essai à partir de leurs expériences respectives dans plusieurs études de cas de partenariats dans l'évaluation. Du point de vue de l'enseignant, si les étudiants deviennent moins dépendants de leurs professeurs, ils peuvent devenir plus responsables de leur apprentissage et autonomes. La négociation implique que l'enseignant ait une position de retrait et soit davantage un guide et un consultant. L'élément de surprise qui ressort des expérimentations est qu'un certain nombre d'étudiants ne prend pas part au changement. Il est possible qu'ils aient internalisé les méthodes traditionnelles d'apprentissage et d'enseignement jusqu'au point où il devient difficile de les faire bouger de leur position. C'est pourquoi le partenariat dans l'évaluation ne convient pas à toutes les situations parce qu'il fait sortir de la zone de confort. Enfin, l'auteure retient que le risque est de donner trop de pouvoir aux étudiants dans la relation enseignant-étudiant.
Du point de vue de l'étudiante, le partenariat est curieux, car les méthodes d'évaluation ne se communiquent pas de manière explicite. Les grilles de notation n'étaient finalement pas différentes des autres modules parce qu'il était difficile d'imaginer les critères d'évaluation différemment. Cependant, les étudiants ont pu exprimer leur point de vue, ce qui a renforcé l'engagement dans la formation. Cela a permis de développer une communauté d'apprentissage, c'est-à-dire une culture commune dédiée à apprendre.

Deeley et Bovill (2015) ont examiné une autre manière de conduire un partenariat évaluatif dans un autre module optionnel de politique publique en première année de master. Le contrôle des connaissances en partenariat se composait d'une grille d'évaluation cocréée par l'enseignant et les étudiants pour deux types d'examens : un dossier de recherche documentaire de 3000 mots qui comptait pour 40 % de la note finale et un examen final de deux heures qui constituait les 60 % restants. La grille de notation faisait l'objet d'une discussion à chaque séance au fur et à mesure de l'avancement des dossiers et du cours. Un questionnaire sur l'activité de création du contrôle des connaissances a été envoyé aux étudiants à la fin du semestre.
Les auteurs ont constaté que les étudiants ont renforcé leur apprentissage approfondi (« *deep learning* ») grâce à l'émergence d'une communauté d'apprentissage : les étudiants avaient un

objectif partagé de conception d'une grille qui allait servir à les évaluer. Leur motivation a été améliorée grâce à la participation active au processus et à la récompense associée à l'engagement, c'est-à-dire la note finale. Il semble que leurs résultats soient meilleurs (sans qu'il soit possible de contrôler les biais de sélection). Cependant, plusieurs étudiants ont été résistants au partenariat, ils n'ont pas souhaité participer au processus : selon eux, le jugement de l'enseignement est supérieur à ce qu'ils peuvent formuler. D'autres ont été seulement déstabilisés par la nouveauté : la réaction initiale était de ne pas comprendre à quoi cela servait de dialoguer sur les critères de notation ; au fur et à mesure des discussions, les étudiants ont été convaincus et ont montré un intérêt à promouvoir leurs propres valeurs et leurs propres jugements sur ce qui devait être pris en compte dans l'évaluation.
Les auteurs concluent que ce genre de dispositif peut modifier la zone de confort des étudiants, surtout s'il entre en confrontation avec ce qu'ils ont internalisé comme norme. Le fait d'énoncer clairement les intentions du partenariat et ses débouchés permet de faire bouger les perceptions des étudiants.

Abdelmalak (2016) a étudié un dispositif partenarial d'évaluation dans un module de formation de master à *l'EDTECH* aux États-Unis. Le chercheur a observé ce cours de quinze semaines, conçu pour initier des enseignants stagiaires ou titulaires aux outils de recherche utilisant des logiciels sur ordinateur. Six étudiants (cinq doctorants et un étudiant de master) âgés de 38 à 55 ans ont été sélectionnés pour participer à des entretiens d'une heure avec le chercheur. Le formateur a d'abord expliqué que le cours serait construit collectivement et que l'évaluation serait négociée. Au cours de la deuxième séance, les étudiants ont réfléchi en groupes pendant trente minutes aux questions posées par l'enseignant pour aboutir à des formes d'exercice commun : est-ce que l'exercice doit être une expérience ? Est-ce que ce doit être un examen écrit ? Une présentation orale ? Un débat ? De ces réflexions, quatre types d'exercices consensuels ont été identifiés : un journal de lecture ; un carnet de recherche ; une évaluation sur les applications étudiées en formation ; un article de recherche. Puis, le formateur et les étudiants ont négocié la manière d'évaluer chaque exercice. Par exemple, pour l'article de recherche, ils ont convenu d'une évaluation par les pairs. L'enseignant et les étudiants sont également convenus d'un barème de notation selon une grille d'évaluation avec plusieurs critères plutôt qu'un simple commentaire écrit. La grille d'évaluation a été proposée dans un premier temps par le formateur puis a évolué avec les idées des étudiants.
Cette manière de procéder a permis d'augmenter le sentiment de contrôle sur la formation, c'est-à-dire que les étudiants ont acquis plus d'autonomie dans l'utilisation de supports d'apprentissage diversifiés, d'accroître la motivation pour apprendre, lorsque les examens choisis conciliaient les besoins personnels, et de concevoir le formateur comme un facilitateur d'apprentissage plutôt

qu'une position surplombante et dirigeant l'acquisition de connaissances. En revanche, les étudiants ont été déstabilisés au début du processus parce qu'ils n'avaient pas l'habitude qu'on leur demande leur avis : ils ont été conditionnés par les pratiques traditionnelles, et il était difficile pour eux de les déplacer de leur zone de confort.

Leslie et Gorman (2016) ont développé un cadre d'étude du partenariat d'évaluation dans une formation d'ingénieurs entre 2011 et 2013 à l'université d'Ashton (Royaume-Uni). L'intervention portait sur l'évaluation du journal de bord des étudiants. Dans le cadre du métier d'ingénieur, le journal de bord revêt une certaine importance parce qu'il permet de garder une trace des activités, de la programmation des projets et il peut être requis pour des considérations légales. Leur étude porte sur deux cohortes de première année : la promotion 2011 contient 96 étudiants, tandis que la promotion 2013 contient 90 étudiants. Pour la première cohorte, les enseignants ont développé leur propre grille de notation, tandis pour la seconde, pour le même module, les étudiants ont participé au processus de redéfinition de la grille, par groupes. Une grille consensuelle a été formulée par les enseignants à partir des retours des étudiants. Cette dernière a été présentée aux étudiants et un questionnaire de satisfaction l'accompagnait pour obtenir leur ressenti sur le processus. Les chercheurs ont comparé les résultats aux examens de mi-semestre et de fin de semestre pour les étudiants de l'année sans intervention et pour ceux de l'année avec l'intervention.
Les résultats sont supérieurs pour la promotion qui a négocié les règles d'évaluation. De plus, les résultats sont en progression sur le semestre. Les étudiants ont majoritairement approuvé le processus de négociation, parce qu'ils ont mieux compris comment le journal de bord était évalué.
Les auteurs notent cependant les limites d'une analyse par cohorte : des caractéristiques d'étudiants différents, des matériaux d'enseignements qui ont été modifiés entre les deux années, la temporalité des évaluations n'étaient pas la même entre les deux cohortes. Ils concluent que malgré ces limites, il existe un réel attrait des enseignants pour modifier l'environnement de l'évaluation.

Doyle, Buckley et Whilan (2017) ont proposé à 159 étudiants de troisième année de licence dans un module de fiscalité de créer leur propre questionnaire à réponses multiples. Ce travail était récompensé par une note qui équivalait à 5 % de la note finale du module. Chaque semaine, un groupe d'étudiants était sélectionné parmi treize groupes du module pour composer cinq questions à choix multiples à partir d'un chapitre d'un manuel de référence qui étaient ensuite posées à l'ensemble des participants. À la fin du semestre, les questions étaient de nouveau

posées dans un partiel qui contenait également des questions créées par les enseignants référents. Les chercheurs ont mesuré le ressenti des étudiants à l'aide d'un questionnaire.
De façon générale, les étudiants ont perçu favorablement ce dispositif, ont mieux compris ce qu'on attendait d'eux dans leurs révisions, et ont apprécié le fait que l'activité de création soit récompensée par une note. En revanche, les étudiants ont pointé un manque de consignes claires dans ce qu'il fallait sélectionner dans le contenu des questions, que certains passages du cours étaient plus « faciles » à préparer que d'autres et qu'ils manquaient de « *feedback* » sur la qualité de leurs questions.
En définitive, les résultats ont été meilleurs à l'examen de fin de semestre qu'à l'examen de mi-semestre, tandis que les questions créées par les étudiants ont été mieux réussies que celles créées par les enseignants.

d. Synthèse de la littérature sur le partenariat évaluation

À la lumière de ces travaux qui ont implémenté la négociation des règles d'examen, que ce soit par le choix de l'évaluation, ou la co-construction de la note, il semble qu'il existe des marges de manœuvre pour modifier le modèle traditionnel du contrôle des connaissances à l'université. Celui-ci pourrait être tourné vers davantage de dialogue entre les enseignants et les étudiants. Ces interventions ont provoqué des effets positifs sur l'engagement des étudiants, leur motivation à l'apprentissage et la prise de responsabilité, tout en créant un environnement favorable à la compréhension de l'évaluation, son intérêt ainsi que ces objectifs. Ces expériences ont montré que le rôle de l'enseignant pouvait être modifié, en passant d'un statut autocratique à celui d'accompagnant disposé à concilier les aspects de transmission de connaissances avec celui de partenaire dans la création de l'évaluation de l'acquisition de connaissances.

Cependant, la portée générale de ces travaux est limitée par plusieurs aspects. Le premier tient aux échantillons d'étudiants considérés sur lesquels portent les interventions. Ces étudiants sont « choisis » dans des contextes plutôt favorables au changement. Par exemple quand il s'agit de module optionnel, les étudiants ont un intérêt pour l'enseignement et sont prêts à accepter des modalités d'évaluation qui puissent différer du reste de leur cursus. La taille de l'échantillon réduit quant à elle la validité externe de certaines interprétations.
À cela s'ajoute le manque d'indicateurs sur la réussite des étudiants. Cette dimension n'est pas majoritairement présente, alors même que les auteurs semblent vouloir innover dans les pratiques d'évaluation dans les formations d'enseignement supérieur pour renforcer la réussite de leurs classes. Même si la majorité des études s'accordent sur le fait que les étudiants sont davantage motivés, la question sous-jacente non résolue est celle de savoir si le gain de motivation se traduit par un gain de performance.

Enfin, parmi ces contributions, aucune n'a conçu de protocole expérimental permettant d'évaluer des effets de causalité des dispositifs sur les étudiants. Les raisons sont multiples : les chercheurs n'avaient pas les moyens de prendre en compte cette méthode dans leur protocole de recherche, par défaut d'acculturation ou par manque d'accord des responsables administratifs.

III. Hypothèses

Dans notre revue de la littérature, nous avons montré que lorsque l'éducation était considérée comme un bien public, les interactions entre les enseignants et les étudiants quant aux modalités d'instruction et d'apprentissage engendrant de la coopération, permettaient d'obtenir une allocation des ressources en effort plus efficace que dans une situation sans négociation. Nous avons également montré que lorsque les deux pôles de la relation avaient des divergences de point de vue sur les bénéfices retirés de l'éducation, c'est-à-dire quand il existe un aléa moral sur le comportement des étudiants ou des enseignants, la négociation et la coopération engendraient la possibilité d'obtenir un contrat d'effort d'apprentissage et d'effort d'instruction plus optimal d'un point de vue social.
Les récentes interventions de chercheurs en science de l'éducation ont tenu compte de la dimension partenariale dans l'enseignement supérieur pour promouvoir des modalités de contrôle des connaissances qui incluent davantage les étudiants au processus de jugement afin de renforcer leur capacité d'auto-évaluation, d'autonomie dans l'apprentissage et de réflexivité sur le niveau de compétences atteint et sur leurs aptitudes.

Même si les études ont montré des effets positifs du partenariat dans l'évaluation, il semble qu'elles soient menées dans des cadres favorables au changement et pour des étudiants plus motivés pour accepter des modifications. Ces mêmes étudiants sont les plus susceptibles de réagir positivement à ces transformations. Il serait donc judicieux de déterminer si ces résultats positifs s'observent sur un échantillon plus diversifié et que l'on pourrait comparer à un échantillon d'individus qui n'en bénéficient pas.
Il serait également intéressant de déterminer si ces alternatives améliorent leur réussite générale à l'université. La question serait donc de savoir si les étudiants qui entrent dans ces partenariats d'évaluation modifient leurs stratégies d'apprentissage dans un module à l'intérieur duquel le partenariat s'effectue, mais également dans des environnements qui contiennent des formes traditionnelles d'évaluation. En d'autres termes, est-ce que l'assimilation et la co-construction des règles d'évaluation pour un module, engendrant un engagement personnel à respecter des règles décidées collectivement, permettraient de favoriser un engagement suffisant pour répondre à des règles qu'ils n'ont pas décidées, mais dont ils comprennent davantage les tenants et les

aboutissants et aux attentes des enseignants dans l'ensemble des modules de formation de l'enseignement supérieur ?

Dans ce chapitre, nous souhaitons ainsi tester l'effet de la négociation des règles d'évaluation dans le cadre d'une relation enseignant-étudiant sur les résultats scolaires aussi bien dans les modules où sont appliquées ces nouvelles modalités de contrôle de connaissance que dans les modules où ne se pratique pas ce dispositif. Puisque la théorie économique prédit des effets positifs sur le bien-être social et que les interventions en recherche-action des sciences de l'éducation préconisent la création de partenariats évaluatifs alors nous formulons les hypothèses suivantes :

Hypothèse 1 : la négociation des règles d'évaluation améliore les résultats scolaires des étudiants dans les modules qui intègrent ce changement grâce à une meilleure assimilation des manières d'évaluer.

Hypothèse 2 : la négociation des règles d'évaluation améliore les résultats scolaires des étudiants de façon générale, grâce à l'adoption de nouvelles normes d'apprentissage pour répondre aux attendus de l'université sur leur acquisition de connaissances et de compétences.

IV. Terrain de recherche

Pour tester ces hypothèses, nous avons réalisé une expérimentation de terrain au sein de deux modules de formation de la licence générale de l'UFR d'Économie-Gestion de l'Université Paris-Est Marne-la-Vallée (Upem).

1. Contexte de l'expérimentation

L'Upem a été créé par le décret n°91-706 du 22 juillet 1991 dans le cadre du plan de modernisation « Université 2000 »[20]. Il dispose de douze composantes de formation, dont cinq instituts (électronique et informatique ; urbanisme ; ingénierie des services ; sciences appliquées ; universitaire de technologie) et sept UFR (ingénieur ; langues et civilisations ; lettres, arts, communications et technologie ; mathématiques ; sciences économiques et de gestion ; sciences humaines et sociales ; sciences et techniques des activités physiques et sportives).

[20] À l'instar des quatre autres « universités nouvelles » d'Île-de-France, ainsi que de deux université bretonnes et lilloises, elle a été créée dans le but d'absorber l'accroissement du flux d'étudiants dans l'enseignement supérieur consécutif de la démocratisation de l'enseignement. Ce plan de construction de nouvelles universités s'est accompagné d'une prise en considération de l'aménagement du territoire afin de contribuer au développement des infrastructures économiques et techniques, des entreprises et des collectivités territoriales.

Lors de l'année 2017-2018, 12 214 étudiants étaient inscrits dans ses formations, soit quatre fois plus que lors de sa création (Gausson, 2018). Le cycle Licence est celui qui accueille le plus d'étudiants (70 % des effectifs) et il se distingue par une importante part d'inscrits en formation professionnelle (22 %). Le taux de boursier (22 %) est inférieur au taux national (35 %), ce qui s'explique par la présence importante d'étudiants apprentis. Les étudiants proviennent majoritairement du bassin de recrutement de l'académie de Créteil (88 %). Enfin, la présence d'étudiants étrangers se situe au même niveau que la proportion nationale (13 %).

Le taux de réussite de l'Upem se situe au même niveau que le taux national : à titre d'exemple, 33 % des néo-bacheliers inscrits en 2011-2012 validaient leur licence en trois ans et 35 % des étudiants ne se réinscrivent pas au bout de la première année (Théophile, 2017). L'Observatoire des formations et des insertions professionnelles des étudiants (Ofipe) calcule également les taux de passage en deuxième année : entre 2007 et 2016, il fluctue autour de 45 %.

a. L'Unité de formation et de recherche d'Économie-Gestion de l'Université Paris-Est Marne-la-Vallée

En ce qui concerne l'UFR d'Économie-Gestion, il accueille chaque année environ un quart des étudiants inscrits à l'Upem, réparti à 71,6 % en licence et 28,4 % en master. La majorité de l'effectif est inscrite dans la licence générale d'Économie-Gestion qui offre un tronc commun de formation les deux premières années et des parcours de spécialisation en économie générale et en gestion la troisième année.

Pour chaque année de licence, ce sont 500 heures de formation qui sont dédiées aux étudiants et la primauté est donnée aux enseignements qui couplent les cours magistraux avec des travaux dirigés. La licence est gérée par une responsable de formation professeure des universités en sciences économiques et chaque année de la licence est sous la responsabilité d'un enseignant-chercheur ou professeur agrégé du secondaire. Les enseignements sont assurés par les enseignants-chercheurs ou les professeurs agrégés tandis que les travaux dirigés sont pris en charge par des doctorants en mission d'enseignements, des attachés temporaires d'enseignement et de recherche, des vacataires et des professeurs agrégés.

En termes de réussite, les étudiants de l'UFR d'Économie-Gestion ont un meilleur taux de passage en deuxième année (55 %) que les étudiants de l'UFR de Sciences et techniques des activités physiques et sportives (STAPS), Mathématiques et Lettres Arts Communication et Technologie (LACT). Il est néanmoins inférieur à celui des Langues. L'UFR d'Économie-Gestion se classe également à la deuxième position du taux de validation de la licence en trois années (34 %). Ces scores démontrent une marge de progression importante pour améliorer le sort des étudiants accueillis dans la licence.

Sur les raisons qui expliqueraient la difficulté des étudiants à réussir dans les formations d'Économie-Gestion, la littérature renvoie aux différents profils de bacheliers qui ne sont pas tous formés au même niveau aux outils mathématiques qui sont utiles pour la capacité d'abstraction, de modélisation et de logique, aussi bien en économie qu'en gestion (Arnold et Straten, 2012). Ainsi, à caractéristiques sociodémographiques équivalentes, un bachelier scientifique aura davantage de réussite qu'un bachelier issu de la filière Science et technologies de management et de gestion (STMG) parce qu'il aura reçu davantage d'heures en mathématiques pour un programme plus technique. La littérature cite également les différents niveaux d'acquisition d'une culture générale qui est mobilisée à divers degrés dans la formation.

Le phénomène de l'échec à l'université est pris en considération par l'équipe dirigeante et l'équipe pédagogique qui a instauré plusieurs dispositifs successifs au cours des années. Parmi les outils de remédiation, on peut en citer au moins quatre qui structurent la formation dès la première année et mettent en condition les étudiants pour leur fournir un cadre d'apprentissage plus immersif.
Le premier est le dispositif de remise à niveau en mathématiques qui permet à des groupes d'étudiants de revoir ou d'approfondir les principales notions enseignées dans le programme du lycée pour être capable d'assimiler celles qui seront transmises à l'université. Lors de la prérentrée, les étudiants réalisent un test de positionnement qui leur autorise à accéder ou non à ce module qui dispose par ailleurs d'un nombre limité de places. Cette forme de tutorat est assurée par l'enseignant référent du module principal de mathématiques et s'effectue durant trois heures hebdomadaires au cours des dix premières semaines de formation du premier semestre de licence. Il permet en outre de valider 3 ECTS. Sans qu'il y ait eu de véritables évaluations du dispositif, il semble néanmoins apporter aux étudiants un accompagnement pour la révision et ils peuvent poser des questions plus facilement qu'en travaux dirigés ou en cours magistral. En revanche, l'une des limites est que certains étudiants qui disposent du niveau requis en mathématiques échouent volontairement au test de positionnement pour être assurés de remporter les crédits de validation du semestre. Par conséquent, ils prennent la place d'étudiants qui auraient le plus besoin de ce dispositif.

Le deuxième outil est un module commun à l'ensemble de la promotion, qui se positionne sur les deux premières années de la licence et qui bénéficie d'un crédit d'heure au titre de l'innovation pédagogique et de l'insertion professionnelle (article 7 de l'arrêté du 22 janvier 2014). La première année, lors de la semaine de prérentrée un jeu d'orientation est organisé sur le campus afin de provoquer des rencontres entre les étudiants de la promotion dans des groupes formés aléatoirement. L'idée étant que ces rencontres créent un premier réseau de fréquentations pour

faire circuler les informations sur l'organisation de la formation, les emplois du temps, les règles de fonctionnement de l'université, etc. À plusieurs moments du semestre, ces groupes se réunissent avec des tuteurs (membres de l'équipe pédagogique ou des étudiants d'années supérieures) pour faire le point sur leurs conditions d'études. La deuxième année, les étudiants formalisent leurs réflexions au sujet de leur insertion professionnelle dans un dossier qui contient leur projet d'étude, leur projection dans un métier et leurs entretiens avec des personnes en activité.

Le fait d'inciter les étudiants à se focaliser sur leurs projections en termes d'insertion professionnelle et sur la constitution d'un réseau de connaissances contribuerait à mieux les intégrer dans la formation et leur ferait réaliser qu'il existe des étapes à franchir après l'obtention du diplôme. De surcroît, ce module aiderait à déterminer si le cursus choisi est le plus opportun pour leur choix de parcours professionnel. Le dialogue dans les groupes permettrait de prendre conscience des questionnements communs que les étudiants formulent afin de les partager et les résoudre collectivement.

Ce module est également récompensé par 3 ECTS. La manière d'évaluer l'implication des étudiants centrée sur la réalisation d'un dossier ou sur des entretiens de fin de semestre a plutôt conduit à des stratégies de mise en conformité à la règle d'évaluation plutôt qu'au développement de raisonnement réflexif sur leur avenir. En d'autres termes, à quelques exceptions près, ce dispositif ne semble pas provoquer de remise en question des étudiants sur leur avenir tant scolaire que professionnel.

Le troisième dispositif est la mise en place d'un suivi particulier de l'assiduité des étudiants pendant les travaux dirigés et son intégration dans les modalités de contrôle des connaissances. Tandis que le suivi de l'assiduité peut être fluctuant selon les sites de formation, l'UFR d'Économie-Gestion dispose d'un outil de gestion informatisé intitulé « Fichier de Gestion des Absences et des Résultats/notes de TD (Figaro) » qui répertorie la présence des étudiants et leurs notes en contrôle continu. Il permet d'inciter les étudiants à être présents à l'université au moins pour assister aux travaux dirigés tandis que la présence aux cours magistraux reste conseillée, mais facultative. Deux règles d'assiduité conduisent à une défaillance : plus de deux absences à un module ; une absence à l'un des deux examens en contrôle continu.

Si le système permet un contrôle coercitif de l'assiduité, il engendre également des comportements d'optimisation du temps. L'absence est vue comme un droit et est utilisée à des périodes les moins opportunes, généralement les semaines avant les examens, pour « prendre le temps de la révision » ou lors de la dernière semaine de formation, pour « profiter d'un moment de répit » alors que ces séances pourraient être un temps d'échange privilégié pour des questions plus précises et pour du « *feedback* ».

Enfin, le dernier dispositif général est le module de méthodologie de travail universitaire. Généralisé dans toutes les licences, il est utilisé de différentes façons par les responsables de formation. À l'UFR d'Économie-Gestion, il est destiné à donner des outils de recherche documentaire et de planification pour la rédaction des documents longs spécifiques à la formation (études de cas d'entreprises, rapports d'études économiques) ainsi que des outils pour la présentation orale. C'est également dans le cadre de ce module qu'est intervenue l'expérimentation du Projet Voltaire (Bellity et al., 2017), une plateforme en ligne dédiée à renforcer l'acquisition des outils de la langue française en mêlant leçons et exercices. Les auteurs de l'expérimentation ont montré que le renforcement en orthographe permettait un gain de points dans les matières littéraires et mathématiques. Depuis 2016, les étudiants réalisent le test de positionnement et effectuent le test final qui rentre dans le calcul de la note de MTU pour valider le module. En plus du projet Voltaire, la dernière version de ce module reposait sur la création d'un document de type article de vulgarisation scientifique pour permettre de développer la compétence de problématisation et de rédaction d'une réponse argumentée autour du thème directeur de l'innovation.

Bien qu'il comporte une dimension d'accompagnement à la réussite, en distillant les outils de méthode pour répondre aux attendus de l'université, le module est dépendant de la façon d'appréhender ce qu'est la méthode universitaire. En l'absence de pratiques pédagogiques communes et partagées, les enseignants disposent de peu de marges de manœuvre pour appliquer des techniques de méthodologie qui permettraient d'améliorer l'effort d'apprentissage des étudiants. Ces derniers considèrent ainsi ce module de formation comme optionnel, en lui attribuant davantage une connotation de remise à niveau de français plutôt qu'un moyen d'acquérir des compétences nouvelles qu'ils apprennent par ailleurs dans les autres modules sur le tas.

b. Les modalités de contrôle des connaissances : une inertie malgré des changements micro-organisationnels

À contrario du développement de dispositifs pédagogiques innovants, les instruments d'examen des connaissances demeurent imperméables au changement. Les règles en la matière sont quasiment communes à l'ensemble de la maquette de la licence.

Lorsqu'un module dispose d'une combinaison cours magistral/travaux dirigés, les étudiants sont examinés au minimum deux fois par semestre et par module au titre du contrôle continu. La plupart du temps,ces évaluations sont soumises à une règle de traitement équitable, c'est-à-dire que, dans la mesure du possible, ces évaluations sont similaires dans leurs formes et leurs contenus. Dans cette configuration, le contrôle continu représente la moitié de la note finale qui

permet de valider l'acquisition d'ECTS. L'autre moitié est représentée par un contrôle terminal, la plupart du temps, sous une forme d'examens partiels effectués par tous les étudiants au même moment à la fin du semestre. Lorsqu'un module s'effectue sous la forme d'un cours magistral, l'examen partiel unique est appliqué.

La forme des examens varie peu et la primauté est donnée aux examens écrits. Schématiquement, les matières d'économie demandent plutôt une capacité de résolution de problèmes de modélisation et de raisonnement faisant appel à la logique et à l'abstraction, tandis que les matières de gestion sollicitent des étudiants leurs capacités d'acquisition d'un ensemble de connaissances générales, d'analyse pour répondre à des questions portant sur un ensemble documentaire et de résolution de cas pratiques. Cela passe donc par des sujets d'examens utilisant des questions appelant des réponses courtes, mais argumentées et justifiées.

Il faut néanmoins nuancer le propos sur l'inertie de l'évaluation parce que des tentatives d'implantation de méthodes alternatives d'évaluation liées à l'auto-évaluation et l'évaluation par les pairs ont émergé.
Cependant, dans plusieurs cas, ces pratiques ont été vécues comme un échec par les enseignants qui ont constaté un manque d'engagement de la part des étudiants dans le processus, voire l'émergence de stratégie de contournement de ces dispositifs. De plus, ces évaluations alternatives n'ont pas été suffisamment étudiées et préparées en amont de leur implémentation ce qui a contribué à une insatisfaction de la part des tuteurs et des étudiants.
Il est possible de noter un manque d'accompagnement des étudiants qui n'ont pas su comment évaluer leur performance ou celle de leurs pairs, n'étant pas coutumiers de tels dispositifs. En définitive, ces pratiques n'ont pas été jugées assez efficaces pour que les scores obtenus à l'issue de ces évaluations soient inclus dans le calcul de la note finale des étudiants. Malgré les bénéfices attendus de ces examens alternatifs (Boud et Falchikov, 2007), ces limites ont affaibli leur portée et leur possible déploiement à d'autres modules de la licence.

Par conséquent, malgré ces apparents changements, l'essence de l'évaluation reste identique et respecte le modèle traditionnel de l'examen dans la relation étudiant-enseignant : l'enseignant conserve l'initiative de l'évaluation et l'étudiant ne dispose pas de marges de manœuvre pour négocier une application alternative de l'évaluation. Dans le cadre des modules où le cours magistral est couplé avec des travaux dirigés, l'enseignant référent adopte une forme d'évaluation qui est ensuite appliquée par les chargés de travaux dirigés.

c. L'évaluation de la participation en tension

Dans cette architecture globale de l'évaluation, nous avons diagnostiqué un point de tension qui provoque une insatisfaction généralisée. En effet, dans le cadre du contrôle continu il existe, dans la majorité des modules, une part attribuée à l'évaluation « générale » des étudiants qui complète le jugement effectué sur leur niveau de connaissances factuelles lors des examens semestriels. Il s'agit de l'évaluation de la participation des étudiants. Selon les choix d'évaluation dans les modules, ce score peut être compris entre 5 % et 15 % de la note finale.

Dans l'UFR d'Économie-Gestion, elle ne bénéficie pourtant pas d'une définition commune et partagée. Elle peut être conçue pour comptabiliser le nombre d'interventions orales des étudiants dans les modules ou le nombre de rendus volontaires d'exercices contenus dans les fascicules de travaux dirigés préparés.

Aux yeux des étudiants, l'ingénierie de l'évaluation de la participation est opaque, c'est-à-dire qu'elle n'est visible qu'en recomptabilisant une moyenne donnée à la fin du semestre. Elle n'influence que rarement le comportement qu'elle est censée instituer à savoir la participation active des étudiants dans les travaux dirigés. Aux yeux des enseignants chargés de travaux dirigés, c'est davantage une case à remplir qu'un outil d'évaluation des connaissances ou des compétences.

Il est pourtant possible de considérer d'une manière différente cette partie de l'évaluation en contrôle continu. Participer signifie collaborer à quelque chose, c'est-à-dire s'associer à une activité dans le but de réaliser un objectif commun. La participation est donc un concept d'action. Celle-ci repose sur des capacités individuelles particulières comme l'autonomie, la prise d'initiative, la réflexion, le jugement critique, l'analyse et la résolution de problème. Par conséquent, considérer la participation c'est finalement considérer l'ensemble des compétences et dispositions que les étudiants seraient en mesure d'acquérir dans une formation d'enseignement supérieur.

De plus, la participation repose sur l'idée de la contribution qui fait écho à la notion de coopération. La coopération peut s'effectuer avec les autres membres du groupe de travaux dirigés ou bien avec l'enseignant. Évaluer la participation, c'est finalement prendre en considération les interactions qui se jouent dans une salle de classe que nous avons décrits dans la relation entre l'enseignant et les étudiants. Elle permet de mettre en avant l'idée que l'étudiant est un acteur à part entière, qu'il joue un rôle dans le processus d'apprentissage et d'instruction et que sa parole prend autant de valeur que celle du tuteur.

In fine, cette notion évoque l'engagement personnel pour une tâche. Cela permet de prendre en compte le positionnement individuel de chaque étudiant dans un processus plus large.

Par conséquent, c'est une dimension plus vaste que la comptabilisation du nombre d'interventions. Il est possible de prendre en compte la façon dont se déploient les interactions entre étudiants, s'il existe de l'entraide ou au contraire de la rivalité. Il est possible de caractériser le fond de ces interventions : est-ce davantage prendre en considération le nombre de réponses justes à une question ? Est-ce prendre en compte le nombre d'erreurs ? Enfin dans cette réflexion, il faut être en mesure de caractériser la pondération à appliquer à chaque critère pris en compte dans chaque plan de la dimension participative. À l'inverse de considérations uniques et uniformes, la pondération et la critérisation peuvent ainsi être évolutives et dépendantes du contexte de chaque groupe de formation dans un module donné.

C'est à l'issue de ce diagnostic que nous avons formulé le protocole expérimental suivant.

2. Protocole expérimental

L'expérimentation s'est déroulée lors du premier semestre de l'année universitaire 2017-2018. Dans cette section nous détaillerons le processus expérimental, la méthode d'assignation des étudiants à la condition expérimentale et le dispositif de collecte d'informations. Le traitement expérimental repose sur l'idée de créer un partenariat évaluatif sur la participation des étudiants au sein des travaux dirigés de deux modules de formation au cours du premier semestre : l'introduction à l'économie en première année ; la microéconomie niveau 2 en deuxième année.

a. Traitement expérimental

La séquence du traitement devait permettre de faire émerger une discussion et une négociation par étape entre les étudiants et l'enseignant sur les modalités de l'évaluation de leur participation dans les travaux dirigés.

Lors de la première séance de travaux dirigés, dédiée généralement à présenter les objectifs, le contenu, les méthodes de travail et le calendrier des examens du contrôle continu, le chargé de travaux dirigés responsable de l'expérimentation explicitait les conditions dans lesquelles s'effectuerait l'évaluation de la participation. La proposition de participation à l'expérimentation était énoncée à chaque groupe de travaux dirigés lors de la première séance en ces termes : « Dans le cadre d'une étude sur l'engagement et la réussite des étudiants, votre groupe de travaux dirigés de XXX a été sélectionné pour expérimenter une nouvelle manière d'évaluer la participation des étudiants. L'objectif de cette expérience est de vous intégrer à la construction de la grille d'évaluation et de la pondération de chaque critère d'évaluation de la participation. Vous serez amenés à proposer vos propres critères d'évaluation et à amender ceux qui sont proposés par

votre chargé de travaux dirigés. Une fois la grille validée, elle constituera la base de votre évaluation de la participation qui compte pour 10 % de la note de contrôle continu, ce qui représente au maximum 2 points sur 20 ». Puis le chargé de travaux dirigés sollicitait l'accord de participer ou non cette expérimentation. Tous les étudiants acceptaient de participer au processus. À la fin de cette introduction, le chargé de travaux dirigés invitait les étudiants à réfléchir aux modalités d'évaluation de la participation sans aucune autre directive, pour qu'elles soient discutées lors de la séance suivante.

Au début de la deuxième séance de travaux dirigés, l'enseignant demandait aux étudiants de donner leurs idées sur ce qu'ils pensaient être une bonne façon de juger la participation. Le pouvoir de décision était donc d'abord dans les mains des étudiants qui avaient pu laisser libre cours à leurs réflexions, qu'elles soient saugrenues ou pas. Ce principe permettait de ne pas créer de jugements préconstruits et de ne pas restreindre la créativité ou l'imagination des groupes. Au bout de cinq minutes, la discussion était close, une liste d'éléments consensuels entre les étudiants était arrêtée et la suite de la séance se déroulait normalement.

Entre la deuxième semaine et la troisième semaine de travaux dirigés, l'enseignant disposait d'un temps de réflexion pour valider les critères débattus par les classes et amender les listes de critères. Les critères supplémentaires permettaient de considérer des caractéristiques de la participation que les étudiants n'avaient pas en tête. Dans cette phase, l'enseignant décidait également de la pondération des critères qui était commune à l'ensemble des travaux dirigés. L'idée était de considérer à égale valeur chaque critère, qu'il soit décidé par les étudiants ou proposé par l'enseignant, c'est-à-dire que chaque critère donnait le même nombre maximal de points. Ensuite pour pouvoir juger de l'attribution des scores, une dimension quantitative était énoncée : si le critère se vérifiait durant les deux tiers des séances de travaux dirigés, le critère donnait le maximum de points (20) ; si le critère se vérifiait durant le tiers des séances de travaux dirigés, le critère donnait la moitié de points (10) ; si le critère ne se vérifiait jamais, le critère ne donnait aucun point (0). La moyenne arithmétique était réalisée pour constituer la note de participation.

Au début de la troisième séance de travaux dirigés, l'enseignant distribuait la liste des critères validés issus de la discussion des étudiants et du choix de l'enseignant ainsi que la méthode de notation des critères. Une nouvelle discussion s'enclenchait pour apporter des modifications sur la liste finale. À l'issue de cette discussion, un vote permettait de valider l'adoption de la liste. *In fine*, les cinq groupes de travaux dirigés tests ont eu une grille d'évaluation qui était spécifique à chaque discussion.

Au cours des séances suivantes, le chargé de travaux dirigés tenait les comptes des interactions qui rentraient dans la grille d'évaluation à l'aide d'un tableur. Il ne faisait plus de mention particulière à la participation, laissant les étudiants libres de se conformer à la règle négociée.

À la fin du semestre, une note était calculée pour chaque étudiant et elle était transmise à l'administration pour le calcul de la note de contrôle en continu. Lors de la dernière séance, la plupart du temps, consacrée au rendu des seconds examens de connaissances, l'enseignant communiquait à chaque étudiant sa grille personnelle qui rappelait les principes de l'expérimentation, mettait en valeur les critères validés et la note finale. Un commentaire général et/ou spécifique à chaque critère était également rédigé afin de donner des recommandations pour la suite du cursus de formation.

b. Assignation aléatoire

Trois modules d'introduction à l'économie et deux modules de microéconomie 2 ont été choisis pour cette expérimentation. Le choix des modules a été décidé d'une part, de manière pragmatique parce qu'ils comptaient parmi le service d'enseignement d'un attaché temporaire d'enseignement et de recherche, d'autre part à partir de préoccupations sur la volonté d'améliorer les résultats des étudiants dans ces modules traditionnellement peu propices à la réussite. De plus, mettre en place le dispositif expérimental dans les groupes d'un seul chargé de travaux dirigés permettait de contrôler les biais pédagogiques propres à chaque tuteur.

Les groupes ont été répartis aléatoirement au chargé de travaux dirigés par le secrétariat pédagogique en fonction de l'emploi du temps personnel de l'enseignant et de l'organisation générale du planning de formation qui dépend des salles disponibles pour les créneaux choisis et du chevauchement d'autres modules aux horaires considérés. Nous sommes donc assurés qu'il n'y a pas eu de sélection particulière dans l'attribution des travaux dirigés selon la liste des étudiants.

De plus, les étudiants sont répartis dans les groupes selon un panachage aléatoire réalisé par le secrétariat pédagogique. Cette répartition dépend à la fois du choix des unités d'enseignements libres (les options) de manière à avoir au sein d'un même groupe de travaux dirigés des étudiants suivant la même option et de l'ordre alphabétique de leur nom de famille. Ainsi, les étudiants ne « choisissent » pas leur emploi du temps. Enfin, après l'initialisation des groupes, aucun changement n'est autorisé par les responsables de formation.

Le groupe témoin est constitué par les cinq autres groupes de travaux dirigés de première année et les quatre groupes de travaux dirigés de deuxième année.
En définitive, on se rapproche d'une expérimentation contrôlée grâce à l'affectation aléatoire dans des groupes répartis aléatoirement dans le planning de formation. Dans le cadre de l'expérimentation du Projet Voltaire (Bellity et al., 2017), ce même processus d'affectation des étudiants aux groupes de travaux dirigés a été assimilé à une assignation aléatoire.

c. Collecte des données

Nous souhaitons évaluer l'effet de ce traitement sur les résultats des étudiants à tous les examens du premier semestre, que ce soit en contrôle terminal ou en contrôle continu, à la fois dans les modules dans lesquels a été conduite l'expérimentation et pour les modules sans modification.

Pour réaliser cette évaluation, nous utilisons une extraction des données issues de l'application pour l'organisation et la gestion des enseignements et des étudiants (Apogée). L'application est structurée en « unités de gestion » des informations issues des inscriptions administratives (état civil, frais d'inscription, sécurité sociale) et pédagogiques (rattachement à un diplôme, une année, à un ensemble de modules de formation), des modalités de contrôle des connaissances et des épreuves (barème, nombre de sessions, règles de calcul, planning), des résultats et des notes (saisie des notes et de la situation d'ajournement ou de défaillance) et de la production des diplômes (procès-verbaux, annexe au diplôme, etc.)[21].

Nous avons réalisé une extraction d'Apogée pour les inscrits au premier semestre 2017-2018 qui contient les noms et prénoms des étudiants, leur date de naissance, et pour chaque module, les notes finales du contrôle continu et du contrôle terminal. Si les étudiants ne respectaient pas les modalités de contrôle des connaissances, les mentions de défaillance étaient renseignées.
Pour les étudiants de première année, nous utilisons leurs résultats du premier semestre de la première session. Il n'existait pas de mention spécifique pour identifier les étudiants qui étaient redoublants dans la promotion. Nous avons donc utilisé les résultats de la promotion 2016-2017 de première année pour identifier les redoublants.
Pour les étudiants de deuxième année, nous procédons de la même façon, à savoir l'édition d'une extraction contenant les résultats de la première session du premier semestre de l'année 2017-2018, ainsi qu'une extraction des résultats du premier semestre de l'année 2016-2017 pour identifier les redoublants.

[21] À cause du nombre de renseignements sensibles et pour éviter des fraudes aux diplômes liées à l'utilisation malhonnête de l'application, les accès à Apogée sont restreints à un faible nombre d'utilisateurs. Nous avons fait appel au secrétariat pédagogique de la licence générale pour avoir accès à ces informations.

Comme nous disposons d'informations de la première année de la promotion 2016-2017, nous pouvons identifier les étudiants qui sont passés de la L1 à la L2 ainsi que les personnes qui proviennent d'une autre formation en réorientation. Pour les premiers, nous disposons ainsi d'informations sur leurs résultats scolaires antérieurs, tandis que pour les seconds, ces informations ne sont pas disponibles.

Nous utilisons également les bases de gestion de l'outil Figaro. Ce dernier permet d'identifier les groupes de travaux dirigés dans lesquels les étudiants sont inscrits, leur présence hebdomadaire et la décomposition des notes du contrôle continu, ainsi que leur pondération pour chaque module.

Enfin, nous avons collecté les informations issues de la plateforme Admission Post-Bac (APB) remplacée au printemps 2018 par la plateforme ParcourSup. APB était conçue pour la saisie des vœux d'inscription des lycéens futurs bacheliers ou des personnes en reprise d'études dans les formations d'enseignement supérieur. La personne responsable de la gestion des candidats a utilisé cette plateforme pour extraire des informations sur la filière d'origine du baccalauréat des étudiants et sur leurs notes obtenues au baccalauréat. Cependant, nous disposons de ces informations uniquement pour les étudiants dont les vœux ont été validés à la fin du processus de sélection. Par conséquent, il n'est pas possible d'obtenir ces informations pour les étudiants qui se sont réorientés grâce à des dispositifs internes à l'université, ou des bacheliers ayant été affectés par le rectorat pendant la phase complémentaire d'assignation. Ces caractéristiques sociodémographiques, au même titre que le lieu de domicile, ne font pas l'objet d'une saisie systématique du secrétariat pédagogique.

d. Comité de suivi et consentement éclairé des étudiants

Comme l'expérimentation de terrain est introduite dans un contexte réel et pas en laboratoire, elle peut bouleverser l'organisation de la formation. Le tirage au sort pour l'assignation au dispositif expérimental pose notamment une question d'ordre morale. En effet, l'attribution aléatoire de la condition expérimentale produit une rupture d'égalité « des chances » ex post en privant les étudiants de ressources temporaires qui pourraient améliorer (détériorer) leur situation. En l'occurrence dans le cadre de la licence, c'est prendre le risque d'affaiblir (d'augmenter) les chances de réussite aux examens et *in fine* de voir les étudiants être pénalisés (favorisés) dans l'obtention de leur diplôme. Cette rupture du principe d'égalité justifie le caractère temporaire d'une expérimentation (L'Horty et Petit, 2011).

C'est pourquoi afin de se prémunir contre d'éventuels biais d'organisation, l'expérimentation a bénéficié de la constitution d'un comité de suivi *ad hoc* composé de la responsable du cycle

licence, du responsable de la première année et de la responsable de la deuxième année. Ce comité a permis d'avoir un regard sur le dispositif et de valider chaque étape du processus pour s'assurer de la pertinence éthique. Les enseignants référents des modules étaient également tenus au courant du dispositif.

À partir des recommandations de ce comité, de la prise en compte du règlement général de la protection des données (RGPD) qui renforce les prérogatives des individus collectant des informations personnelles, ainsi que des préconisations en termes de démarche scientifique du Centre national pour la recherche scientifique (CNRS)[22], nous avons demandé le consentement des étudiants quant à l'utilisation de leurs informations personnelles dans le cadre du dispositif expérimental.

Ainsi, les promotions de première année et de deuxième année ont signé un accord de consentement de l'utilisation de leurs informations dans lequel était mentionné une description du cadre du projet de recherche et de la finalité de la collecte, que les données collectées portaient sur les résultats scolaires ainsi que sur d'autres informations administratives, et qu'il était possible de consulter ou de supprimer l'accès à ces informations.

Dans le cadre de cette démarche, nous avons pu avoir le consentement de 70 % des étudiants des deux promotions. En revanche, une des limites de cette décision est que plusieurs étudiants des groupes traités ont refusé l'accès à leurs informations personnelles. En respectant leur volonté, nous n'avons donc pas utilisé leurs résultats dans le cadre de notre étude. Cela peut donc constituer un biais dans nos interprétations finales.

V. Données

1. Étude de l'échantillon

Lors du premier semestre de l'année 2017-2018, la promotion de première année de licence générale d'Économie-Gestion comptait 289 inscrits et la promotion de deuxième année comptait 202 inscrits.

Parmi ces inscrits, nous enlevons les étudiants qui relèvent de la situation « prise en compte pour calcul » (Prc) : ce sont des étudiants qui sont inscrits pédagogiquement pour l'année en cours, mais qui ont validé l'un des deux semestres l'an passé. Dans ce cas, ils redoublent le semestre qu'ils n'ont pas validé et peuvent bénéficier d'un passage dans l'année supérieure. Il existe donc des étudiants qui ne sont pas présents dans les travaux dirigés du premier semestre. Par

[22] Notamment le guide « Pratiquer une recherche intègre et responsable » du Comité d'éthique du CNRS (Comets) qui préconise la protection des données à caractère personnel à travers cinq principales recommandations : la sécurisation des fichiers ; la confidentialité des données ; le principe de finalité pour indiquer le but de la collecte ; le principe de proportionnalité qui permet de fixer la quantité et la durée de la collecte ; l'information des sujets d'étude.

conséquent, nous les retirons de l'analyse pour ne conserver que les étudiants qui sont en travaux dirigés puisque le traitement s'applique dans ce contexte.

De plus, comme nous l'avons expliqué dans la partie de présentation du protocole expérimental, nous retirons également de l'analyse les étudiants qui n'ont pas consenti à donner un accès à leurs informations personnelles. Les caractéristiques des étudiants du groupe traité et du groupe témoin sont renseignées dans le tableau 1 et le tableau 2.

Tableau 1 : Statistiques descriptives et test d'équivalence entre le groupe traité et le groupe témoin des étudiants de première année de licence d'Économie-Gestion 2017-2018

	Groupe traité	Groupe témoin	Différence	Test de comparaison	
				Test de Wald	Test de Student
Part de femmes	45,455 (5,674)	45,283 (4,835)	-0,172 (7,454)	-0,023	
Part de redoublants	11,688 (3,661)	10,377 (2,962)	-1,311 (4,672)	-0,2806	
Part de réorientation	19,481 (4,513)	19,811 (3,871)	0,331 (6,188)	0,0556	
Part de néo-entrants	68,831 (5,278)	69,811 (4,459)	0,98 (6,900)	0,142	
Taux de BAC ES	62,338 (5,522)	63,208 (4,684)	0,87 (7,236)	0,1202	
Taux de BAC S	16,883 (4,269)	16,981 (3,647)	0,098 (5,617)	0,0175	
Taux de BAC autre (L, STMG, P)	11,688 (3,661)	11,321 (3,077)	-0,368 (4,772)	-0,077	
Âge (au début du traitement)	19,221 (0,996)	19,120 (1,158)	-0,101 (0,163)		-0,6165
Moyenne générale au baccalauréat	11,949 (1,438)	12,014 (1,591)	-0,065 (0,280)		0,2329
Nombre de groupes de travaux dirigés	3	5			
Nombre d'étudiants	77	106			

Note : Les données entre parenthèses sont les erreurs-types pour les variables discrètes et les écarts-types pour les variables continues.

Il y a davantage d'étudiants inscrits que d'étudiantes dans la formation. La part de redoublant du premier semestre concerne un étudiant sur dix. Deux étudiants sur dix n'étaient pas en terminale l'année précédente. La majorité de la promotion est diplômée d'un baccalauréat économique et social. L'âge moyen au début de l'expérimentation est de 19 ans, ce qui correspond à un an de retard avec l'âge « normal » des étudiants à l'entrée de la licence. Cela reflète bien l'idée que la promotion contient des étudiants redoublants ou en réorientation. La moyenne générale obtenue au baccalauréat est de 12/20, ce qui équivaut à la mention « assez bien », et la promotion ne semble pas spécialement hétérogène sur son niveau antérieur.

Nous notons que la répartition correspond à une allocation de 40 % des étudiants dans le groupe traité et 60 % des étudiants dans le groupe témoin. Les tests pratiqués sur les caractéristiques

observables dans les deux groupes nous permettent de ne pas rejeter l'hypothèse de nullité des écarts sur les caractéristiques observables au seuil de 5 %.

Tableau 2 : Statistiques descriptives et test d'équivalence entre le groupe traité et le groupe témoin des étudiants de deuxième année de licence d'Économie-Gestion 2017-2018

	Groupe traité	Groupe témoin	Différence	Test de comparaison	
				Test de Wald	Test de Student
Part de femmes	45,565 (7,344)	41,111 (5,187)	-4,454 (8,990)	-0,5066	
Part de redoublants	4,348 (3,007)	17,778 (4,030)	**13,43**** **(5,028)**	**-2,187**	
Part de réorientation	6,522 (3,640)	10,000 (3,162)	3,478 (4,822)	0,6766	
Part de néo-entrants	89,130 (4,589)	72,222 (4,721)	**-16,908**** **(6,582)**	**-2,2499**	
Taux de BAC ES	34,783 (7,022)	34,444 (0,501)	-0,338 (8,619)	-0,0392	
Taux de BAC S	21,739 (6,082)	18,889 (4,126)	-2,85 (0,735)	-0,3942	
Taux de BAC autre (L, STMG, P)	4,348 (3,007)	2,222 (1,554)	-2,126 (3,385)	-0,6941	
Âge	20,140 (0,953)	20,084 (0,910)	-0,055 (0,168)		-0,3294
Moyenne générale au Baccalauréat	12,580 (1,582)	12,281 (1,314)	-0,299 (0,350)		-1,1368
Nombre de groupes de travaux dirigés	2	4			
Nombre d'étudiants	46	90			

Note : Les données entre parenthèses sont les erreurs-types pour les variables discrètes et les écarts-types pour les variables continues. ** : p-value < 0,05.

En ce qui concerne la promotion de deuxième année, nous observons qu'il y a également davantage d'étudiants que d'étudiantes. La part de redoublants atteint 13 %, tandis que le taux d'étudiants en réorientation est de 8 %. Lorsque la filière du baccalauréat est renseignée, le mode est la filière économique et sociale. L'âge moyen est de 20 ans, ce qui correspond de nouveau au phénomène de parcours scolaire avec retard. Lorsque la moyenne du baccalauréat était renseignée, les étudiants avaient une note générale de 12/20.

Nous notons que la répartition correspond à une allocation de 30 % de la promotion dans le groupe traité et 70 % dans le groupe témoin. À contrario de la promotion de première année, le groupe témoin est différent du groupe traité sur la répartition des parcours antérieurs des étudiants. Il y a significativement moins de redoublants dans le groupe traité. C'est donc une caractéristique qu'il faudra tenir compte dans les analyses suivantes.

2. Les indicateurs d'intérêt

Notre variable d'intérêt concerne la réussite académique. Nous approchons cette dimension à l'aide de deux indicateurs.

Le premier indicateur considère les résultats semestriels du module dans lequel est intervenu le processus de négociation des règles d'évaluation de la participation aux travaux dirigés. Trois notes sont retenues dans le cadre de l'expérimentation : le premier examen en contrôle continu ; le second examen en contrôle continu ; l'examen partiel en contrôle terminal. Nous nous restreignons ainsi à l'effet de la condition expérimentale sur la réussite dans le module, à diverses étapes de l'effort d'apprentissage des étudiants.

Le second indicateur est une mesure plus globale de la réussite qui prend en compte le fait d'avoir validé ou non le semestre lors de la première session d'examen. Il s'agit dans ce cas d'évaluer l'impact de la condition expérimentale sur la stratégie générale des étudiants et de déterminer si ce dispositif ponctuel pour un seul module a permis de modifier les dispositions des étudiants pour leur stratégie de réussite à l'université.

3. Méthode d'identification des effets

a. L'effet moyen du traitement

L'évaluation causale des effets est permise grâce au modèle de Rubin (1974). Dans ce cadre, nous cherchons à estimer l'effet moyen du traitement, la négociation de la règle d'évaluation du partenariat, sur les notes obtenues lors du contrôle continu et lors du contrôle terminal.

Dans un premier temps, nous comparons les écarts de notes moyennes obtenues aux trois examens du module par le groupe traité et le groupe témoin. Nous réalisons un test d'hypothèse de nullité des écarts observés en utilisant la statistique de Student.

Dans un second temps, nous utilisons les résultats obtenus dans les autres modules afin de contrôler de l'hétérogénéité de la difficulté des autres composantes de la formation. Dans ce contexte on peut alors se rapprocher d'un modèle d'estimation en double différence :

$$Y_{i_m} = \alpha_1 . G_i + \alpha_2 . T_m + \alpha_3 . (G_i . T_m) + \beta + \epsilon_{i_m}$$

où Y_{i_m} correspond à la note d'un étudiant i à un examen (de contrôle continu ou terminal) d'un module m. La variable G_i est une indicatrice d'appartenance au groupe traité ($G_i = 1$) ou au groupe témoin ($G_i = 0$). La variable T_m correspond à l'indicatrice du module traité. La variable d'interaction $G_i . T_m$ correspond à l'appartenance au groupe traité dans le module traité. Par conséquent, α_3 mesure l'indicateur de double différence suivant :

$$\alpha_3 = \left(\overline{Y_{1_1}} - \overline{Y_{0_1}}\right) - \left(\overline{Y_{1_0}} - \overline{Y_{0_0}}\right)$$

où la première différence correspond à l'écart de notes observé uniquement pour le module traité et la seconde différence à l'écart de notes observé pour les autres modules. De cette manière, nous tenons compte du résultat des étudiants dans les autres modules et cela nous permet de capturer un effet pur du traitement dans le module. Si $\alpha_3 > 0$, cela indique que les étudiants du groupe traité ont eu de meilleurs résultats dans le module traité par rapport au groupe témoin et

leurs résultats dans les autres modules. Nous pourrons ainsi attribuer à l'environnement de négociation l'effet causal de cet effet.

Les autres coefficients du modèle mesurent d'autres indicateurs d'intérêt :

$$\beta = \overline{Y_{0_0}}$$

qui correspond à la note moyenne obtenue par le groupe témoin dans les modules non traités ;

$$\alpha_1 = \overline{Y_{1_0}} - \overline{Y_{0_0}}$$

qui identifie l'écart de notes moyennes entre le groupe traité et le groupe témoin pour les modules de formation non traités ;

$$\alpha_2 = \overline{Y_{0_1}} - \overline{Y_{0_0}}$$

qui mesure la différence de notes moyennes obtenues pour le groupe témoin entre le module traité et les modules non traités.

Nous estimons ce modèle par la méthode des moindres carrés ordinaires.

b. Modèle dichotomique pour mesurer l'effet sur la validation du diplôme

Pour estimer l'effet du traitement sur le fait de valider le semestre, nous recourons à un modèle dichotomique de type logit. Il s'écrit de la manière suivante :

$$logit\ P(Y_i = 1|X_i) = \alpha_1 . G_i + \beta_1 . X_i + \alpha_0$$

où Y_i est l'indicatrice de validation du semestre qui vaut $Y_i = 1$ lorsque l'étudiant est admis, et $Y_i = 0$ s'il ne l'est pas. Nous ne rentrons pas dans des considérations de modélisation polytomiques, qui permettraient de tenir compte des résultats en termes de défaillance, d'absences justifiées ou d'étudiants ajournés.

La validation du semestre dépend des caractéristiques observables des étudiants qui rentrent dans la variable X_i : le type de baccalauréat, le genre, le fait d'avoir redoublé. La variable G_i permet d'identifier l'appartenance au groupe témoin. Dans ce cadre, α_1 estime l'effet du traitement sur la réussite. Nous estimons ce modèle grâce à la méthode du maximum de vraisemblance.

VI. Résultats

Dans cette partie, nous présentons les résultats de l'expérimentation sur les indicateurs de réussite scolaire. Nous illustrons une forme de grille d'évaluation de la participation à l'issue de la négociation en annexe (voir annexe 2) et nous réalisons un test d'équivalence sur les notes obtenues sur la participation par le groupe traité (voir annexe 3).

Nous effectuons les évaluations d'impact en distinguant les deux promotions, de première année et de deuxième année, afin de tenir compte de la spécificité de chaque période de formation. De plus, nous tenons compte des résultats à l'issue de la première session d'examen, ce qui implique que les résultats finaux de la deuxième session ne sont pas considérés dans ces analyses. La session de rattrapages se réalise en effet à la fin de l'année étudiante, après les examens du second semestre. Par conséquent, les étudiants réalisent des arbitrages différents lors de cette période, pour maximiser les chances de valider une unité d'enseignements. Il serait possible de mesurer l'effet de durabilité de l'expérimentation sur les étudiants au rattrapage, mais le faible nombre d'observations réduirait le pouvoir statistique de mesure.

Dans un premier temps, nous évaluons l'impact de l'expérimentation sur les résultats aux trois examens de connaissances, d'une part sur le module traité, sans tenir compte des résultats des étudiants dans les autres matières, ce qui permet de déterminer l'effet moyen du traitement, et d'autre part en tenant compte des résultats dans les autres matières, ce qui nous permet de contrôler les stratégies de réussite des étudiants dans l'ensemble de la formation. Dans un second temps, nous évaluons l'impact de l'expérimentation sur les chances de validation du premier semestre pour les étudiants de première année et du troisième semestre pour les étudiants de deuxième année.

1. La réussite aux examens de contrôle continu et terminal

a. Les résultats d'examens des étudiants de première année de licence

En ce qui concerne les étudiants de la promotion de première année de licence d'Économie-Gestion de l'année 2017-2018, le premier examen en contrôle continu du module d'introduction à l'analyse économique était constitué d'un QCM qui portait sur l'acquisition de connaissances générales en économie à partir d'ouvrages de la collection *Repères* lus et étudiés chaque semaine par les étudiants. Le deuxième examen était formé de la même manière et portait sur un ensemble plus important de lectures. Le contrôle terminal contenait un QCM ainsi qu'un exercice de résumé de texte. Cette méthode de travail avait été apprise en travaux dirigés.

Dans les examens en contrôle continu, nous tenons compte des matières de mathématiques, d'anglais, d'introduction à la gestion et de méthodologie de travail universitaire. Pour toutes ces matières, les examens étaient des tests écrits. Pour les examens en contrôle terminal, nous ajoutons les résultats dans les modules optionnels (langue vivante 2, introduction à la sociologie, histoire de la pensée économique) et dans le module de construction de projet d'étude.

Les résultats de l'expérimentation sur la réussite aux examens des étudiants de première année sont inscrits dans le tableau 3.

Tableau 3 : Évaluation d'impact de la condition expérimentale sur les résultats des étudiants de première année

	Résultat CC1	Résultat CC1	Résultat CC2	Résultat CC2	Résultat CT	Résultat CT
Différence simple traité — témoin sur le module traité	**1,043*** (0,389)**		**1,787*** (0,439)**		0,468 (0,421)	
Différence simple traité — témoin sur les modules non traités		-0,469 (0,339)		**0,814** (0,335)**		0,603** (0,282)
Différence pour le groupe témoin entre le module traité et ceux non traités		**-0,945*** (0,344)**		0,165 (0,368)		**-4,608*** (0,350)**
Double différence		**1,512*** (0,516)**		**0,973* (0,551)**		-0,134 (0,505)
Référence : Groupe témoin	10,08*** (0,273)	11,03*** (0,209)	12,70*** (0,297)	12,54*** (0,219)	6,731*** (0,296)	11,34*** (0,189)
Nombre d'observations	174	846	171	831	166	1138

Note : les erreurs-types sont entre parenthèses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Nous constatons un effet positif et significatif de la condition expérimentale, c'est-à-dire le fait d'avoir participé à un processus de co-construction de l'évaluation de la participation dans le module d'introduction à l'économie, sur les résultats du module, à la fois lors du premier examen en contrôle continu et lors du deuxième examen en contrôle continu (colonne 1 et colonne 3). Nous observons même une amélioration des résultats entre les deux examens, plus importante pour le groupe traité, mais cette progression des résultats n'est pas significative.

Lorsque l'on tient compte de la performance des étudiants dans les examens des matières qui n'ont pas disposé de la négociation de la règle d'évaluation de la participation, l'effet positif se maintient (colonne 2 et 4). Cependant, l'écart de réussite consécutif au traitement dans le module traité s'atténue entre les deux périodes d'examen. Nous mesurons par ailleurs que les étudiants du groupe de traitement ont été significativement meilleurs dans les modules non traités que le groupe témoin au cours du deuxième examen en contrôle continu. Enfin, les étudiants traités par le dispositif expérimental ont eu plus de réussite dans le module traité que dans les autres modules. Cet écart de réussite est deux fois plus important lors du second examen que lors du premier examen en contrôle continu, et il est significatif. Cependant, la progression de l'écart de réussite entre les deux périodes d'examen et entre les deux types de modules pour le groupe traité n'est pas significative.

Par conséquent, l'application d'un environnement favorable à la négociation de la règle de l'évaluation de la participation dans les travaux dirigés a permis aux étudiants de s'impliquer

davantage dans le module ce qui s'est traduit par de meilleurs résultats que pour les étudiants qui n'en bénéficiaient pas.

Néanmoins, ce constat positif disparaît lorsque nous considérons les résultats lors de l'examen partiel en contrôle terminal (colonne 5 et 6). Alors que le contrôle terminal se soldait par une note moyenne de 5 points inférieurs à la note de contrôle continu, il ne semble pas que le traitement ait contribué à significativement atténuer ces résultats en baisse pour les étudiants participant à l'expérimentation. Cette absence d'effet s'observe également lorsque nous tenons compte des résultats dans les autres modules de formation. Enfin, nous observons que le groupe traité a eu de meilleurs résultats en moyenne au contrôle terminal pour les modules où le traitement était absent.

b. Les résultats d'examens des étudiants de deuxième année

En ce qui concerne les étudiants de la promotion de deuxième année de licence d'Économie-Gestion de l'année 2017-2018, le premier examen en contrôle continu du module de microéconomie niveau 2 était constitué de deux exercices de modélisation de la théorie du consommateur et du producteur et de questions de mesure des connaissances acquises à partir du cours magistral. Le deuxième examen était formé de la même manière et portait sur des exercices de modélisation de la théorie de l'équilibre général. Le contrôle terminal contenait un QCM ainsi qu'un exercice de modélisation de la concurrence parfaite et imparfaite.

Dans les examens en contrôle continu, nous tenons compte des matières de macroéconomie 2, de mathématiques, d'anglais, de comptabilité et de gestion des ressources humaines. Pour toutes ces matières, les examens étaient aussi des tests écrits. Pour les examens en contrôle terminal, nous ajoutons les résultats dans les modules optionnels (langue vivante 2, sociologie du consommateur, problèmes économiques contemporains) et dans le module de construction de projet professionnel, et de droit des contrats.

Les résultats sur la réussite aux examens des étudiants de deuxième année sont inscrits dans le **tableau** 4.

Tableau 4 : Évaluation d'impact de la condition expérimentale sur les résultats des étudiants de deuxième année

	Résultat CC1	Résultat CC1	Résultat CC2	Résultat CC2	Résultat CT	Résultat CT
Différence simple traité — témoin sur le module traité	-1,137 (0,687)		**-1,845***** **(0,639)**		1,148* (0,634)	
Différence simple traité — témoin sur les modules non traités		0,527 (0,338)		0,0703 (0,294)		**1,274***** **(0,247)**
Différence pour le groupe témoin entre le module traité et ceux non traités		**-2,117***** **(0,506)**		**-4,123***** **(0,471)**		**-1,920***** **(0,426)**
Double différence		**-1,665**** **(0,762)**		**-1,915***** **(0,701)**		-0,126 (0,677)
Référence : Groupe témoin	9,982*** (0,465)	12,10*** (0,204)	8,789*** (0,443)	12,91*** (0,167)	7,963*** (0,403)	9,884*** (0,145)
Nombre d'observations	128	771	128	771	127	1263

Note : les erreurs-types sont entre parenthèses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

L'évaluation de l'impact de la condition expérimentale sur les scores des étudiants de deuxième année conduit à des résultats opposés à ceux de première année. En effet, les étudiants du groupe traité ont moins bien réussi que le groupe témoin aux deux examens en contrôle continu, et ce résultat négatif est significatif lors du deuxième contrôle des connaissances (colonne 3). De plus, alors que les résultats diminuaient pour les deux groupes entre le premier et le deuxième examen en contrôle continu, la baisse des résultats était plus importante pour le groupe traité. En revanche, cet écart n'est pas significatif.

L'écart négatif des résultats entre le groupe traité et le groupe témoin dans le module qui bénéficiait d'un environnement plus favorable à la négociation de l'évaluation de la participation s'observe de nouveau lorsque nous tenons compte des résultats des étudiants dans les autres matières (colonne 2 et 4). En effet, alors que les deux groupes ont des résultats similaires dans les modules de travaux dirigés sans intervention expérimentale, les étudiants traités ont eu significativement de moins bons résultats en microéconomie 2, au cours des deux examens en contrôle continu. Cet effet s'accroît entre les deux contrôles de connaissances. Nous notons également que les étudiants obtiennent un écart de -4 points en microéconomie par rapport aux autres modules.

Par conséquent, l'adoption d'une pratique de négociation de l'évaluation de la participation n'a pas permis d'améliorer les résultats des étudiants. Au contraire cela a davantage renforcé leur échec en diminuant leur réussite aux examens.

En revanche, l'effet de la condition expérimentale sur le résultat des étudiants pendant le contrôle terminal est positif et significatif. Les étudiants du groupe traité ont ainsi obtenu en moyenne 1,2

points supplémentaires au partiel de microéconomie, ce qui a permis de compenser en partie les résultats obtenus dans le cadre du contrôle continu. Cependant, cet écart positif ne s'observe pas quand nous tenons compte des résultats dans les autres matières. Les étudiants sous la condition expérimentale ont ainsi mieux réussi dans les modules non traités que le groupe témoin.

L'hétérogénéité de l'impact du traitement sur la réussite des étudiants interpelle quant à la question de savoir s'il est possible de généraliser le dispositif. Nous y reviendrons dans la partie d'interprétation de cette évaluation d'impact.

2. La réussite en termes de validation du semestre

Dans cette section, nous estimons l'impact de la négociation de l'évaluation de la participation en travaux dirigés sur un indicateur de réussite plus global. En effet, nous souhaitons déterminer si les étudiants qui ont bénéficié de ce dispositif ont conduit des stratégies de réussite qui leur ont permis de valider leur semestre lors de la première session d'examen. Dans ce cadre, les notes seules ne jouent pas sur l'admission, la pondération des modules et les règles de compensation rentrent aussi en considération dans l'effort d'apprentissage. Ainsi, nous évaluons la capacité des étudiants à mieux tenir compte de cet environnement grâce au dispositif qui leur permettait d'acquérir des dispositions d'autonomie et de prise de décision.

Pour cela, nous utilisons le critère de validation du premier semestre, qui est une variable dichotomique, et nous estimons l'effet du traitement sur la probabilité de valider le semestre. Nous ajoutons également des variables de contrôle pour identifier d'éventuels effets de composition : l'âge, le genre, le type de baccalauréat et la situation de l'étudiant l'année précédente.

a. La validation du premier semestre des étudiants de première année

Les résultats de l'estimation du modèle logit pour la réussite des étudiants de première année sont inscrits dans le tableau 5.

Tableau 5 : Estimation de l'effet du traitement sur la probabilité de valider le premier semestre pour les étudiants de première année

	(1)	(2)	(3)	(4)	(5)
Condition expérimentale	0,0798	0,0879	0,0864	0,0857	0,0850
	(0,0732)	(0,0721)	(0,0702)	(0,0703)	(0,0669)
Âge		**-0,0792****	-0,0652	-0,0514	-0,0444
		(0,0390)	(0,0409)	(0,0633)	(0,0575)
Sexe			**0,223*****	**0,225*****	**0,194*****
			(0,0699)	**(0,0703)**	**(0,0693)**
Référence : Néo-entrants					
Redoublants				-0,0558	-0,0232
				(0,165)	(0,156)
Réorientation				-0,0262	-0,0531
				(0,112)	(0,132)
Référence : Baccalauréat E,S					
Baccalauréat S					**0,247****
					(0,0982)
Autre baccalauréat					**-0,300****
					(0,118)
Baccalauréat non-communiqué					0,0797
					(0,135)
Constante	0,227	6,733**	5,452	4,306	4,061
	(0,196)	(3,374)	(3,670)	(5,424)	(5,393)
Pseudo R^2	0,005	0,028	0,067	0,068	0,137
Critère AIC	250,5	246,8	239,1	243,0	231,8
Nombre d'observations	183	183	183	183	183

Note : les effets moyens marginaux sont inscrits pour chaque variable. Les erreurs-types sont entre parenthèses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Nous constatons un effet positif, mais non significatif de la condition expérimentale sur la probabilité de valider le semestre. Cet effet se maintient avec l'inclusion de variable de contrôle. Par conséquent, nous concluons que les étudiants ne semblent pas avoir eu de stratégie d'optimisation des règles de validation lors du premier semestre de la première année de licence lorsqu'ils ont bénéficié d'un environnement qui leur aurait permis de mieux comprendre l'ingénierie de l'évaluation.

Parmi ces variables, nous notons que les femmes ont davantage de chance de valider leur premier semestre que les hommes, de 20 points de pourcentage de différence. Les bacheliers de la filière scientifique montrent également plus de chances de valider leur premier semestre que les bacheliers de la filière économique et sociale.

b. La validation du troisième semestre des étudiants de deuxième année

L'évaluation de l'impact de la condition expérimentale sur la probabilité de valider le troisième semestre de la licence pour les étudiants en deuxième année est inscrite dans le tableau 6.

Tableau 6 : Estimation de l'effet du traitement sur la probabilité de valider le troisième semestre pour les étudiants de deuxième année

	(1)	(2)	(3)	(4)	(5)
Condition expérimentale	**0,140***	**0,148****	**0,140***	**0,148****	**0,162****
	(0,0847)	**(0,0741)**	**(0,0721)**	**(0,0748)**	**(0,0737)**
Âge		**-0,164*****	**-0,162*****	**-0,170*****	**-0,150*****
		(0,0424)	**(0,0411)**	**(0,0486)**	**(0,0533)**
Sexe			**0,194*****	**0,196*****	**0,204*****
			(0,0746)	**(0,0744)**	**(0,0747)**
Référence : Néo-entrants					
Redoublants				0,0465	0,0824
				(0,116)	(0,117)
Réorientation				0,0234	0,0792
				(0,127)	(0,132)
Référence : Baccalauréat E,S					
Baccalauréat S					**-0,182***
					(0,105)
Autre baccalauréat					-0,319
					(0,214)
Baccalauréat non-communiqué					-0,165
					(0,103)
Constante	0,314	16,51***	16,68***	17,42***	16,50**
	(0,214)	(5,152)	(5,415)	(6,106)	(6,571)
Pseudo R^2	0,014	0,100	0,135	0,136	0,163
Critère AIC	181,4	167,9	163,7	167,5	168,6
Nombre d'observations	136	136	136	136	136

Note : les effets moyens marginaux sont inscrits pour chaque variable. Les erreurs-types sont entre parenthèses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Nous observons que pour les étudiants qui sont entrés dans un partenariat évaluatif dans le module de microéconomie, leur chance de valider leur troisième semestre augmente de 14 points de pourcentage. Cet effet positif se maintient avec l'inclusion de variables de contrôle. Par conséquent, bien que leurs notes en microéconomie aient été significativement inférieures à celles du groupe témoin, les étudiants ont adopté des stratégies d'optimisation de leurs résultats dans les autres matières. Il est possible d'envisager que les étudiants aient intégré des dispositions de prise en compte des règles d'évaluation pour maximiser leurs chances d'obtenir de meilleurs résultats dans les modules sans environnement de collaboration sur l'implantation de l'évaluation de la participation.

Pour la promotion de deuxième année, on constate que plus les étudiants sont âgés, moins ils ont de chance de valider leur semestre. Le fait d'être une femme augmente la probabilité de valider le semestre de 20 points de pourcentage. De manière surprenante, les étudiants passés par la filière scientifique du baccalauréat ont moins de chance de valider leur semestre lors de la première session. Néanmoins, l'interprétation de ce coefficient est à prendre avec précaution, parce que l'information sur le diplôme du baccalauréat était mal renseignée pour cette promotion.

VII. Discussion et interprétations des résultats

À l'issue de notre expérimentation de la négociation des règles d'évaluation de la participation dans les travaux dirigés d'introduction à l'économie en première année et de microéconomie en deuxième année de licence d'Économie-Gestion, nous obtenons des résultats hétérogènes et des effets qui semblent opposés : un gain de performance pour les uns ; une détérioration de performance pour les autres.

1. L'ambivalence du traitement pour les étudiants de première année

En ce qui concerne les étudiants de première année, le dispositif expérimental nouant une interaction entre l'enseignant et les étudiants pour coopérer sur les règles d'évaluation semble permettre aux étudiants d'acquérir des capacités d'autonomie et de prise de décision dans leur apprentissage, ce qui a engendré une réussite supérieure lors des deux examens en contrôle continu pour le module traité et lors des seconds examens en contrôle continu pour les modules bénéficiant d'une condition expérimentale et pour ceux n'en disposant pas.

Par conséquent, les étudiants ont réagi favorablement à ce dispositif et ont manifesté un effort dédié à se conformer aux règles coconstruites de la participation au module d'introduction à l'économie, effort qui s'est également traduit par de meilleurs résultats aux examens de contrôle continu. Dans ce contexte, les étudiants ont adopté un cadre d'apprentissage favorable à la réussite en développant dans ces groupes une communauté d'apprentissage, à l'instar de ce qui était observé dans les interventions de chercheurs en science de l'éducation (Deeley, 2014). Cela s'est traduit par la visibilité d'interrelations entre les étudiants pendant les travaux dirigés en introduction à l'économie, qui se sont sans doute reproduites dans les autres travaux dirigés. Si l'on considère que ces échanges entre étudiants existent également dans les autres groupes, ils ont été davantage sujets à faire progresser les étudiants pour assimiler les connaissances et les règles d'organisation de la licence.

L'acceptation de cette intervention est renforcée par l'idée que les étudiants font, dans leur majorité, leurs premiers pas à l'université. Ils n'éprouvent donc pas encore tout le fonctionnement sous-jacent à l'organisation de la licence : le contrôle de l'assiduité, la gestion de l'emploi du temps, la position particulière des enseignants (enseignant-chercheurs, vacataires, doctorants, professeurs agrégés…), le type de connaissances à assimiler. Ils sont donc en condition d'absorber le discours des enseignants pour en saisir les traits de fonctionnement qu'ils leur permettront de réussir et diminuer la peur de l'échec qui demeure ancrée dans leurs cadres de pensée. Dans la condition expérimentale, ils sont ainsi en mesure de « jouer le jeu » puisqu'il n'y a pas d'autres conséquences que la note de participation sur laquelle ils ne savent pas encore dans quelle mesure elle aura un impact sur leur réussite. Les dispositions qu'ils développeront dans ce contexte seront alors jugées valables pour leur cursus à la fois pour le module dans lequel se déroule l'expérimentation et pour les autres modules de formation. La mise en pratique de ces dispositions dans un contexte non expérimental a permis d'engranger une réussite supérieure aux autres étudiants.

La conséquence de cette réception favorable a été un effort important pour atteindre de hauts scores pour le contrôle continu. Cependant, lorsque la période des examens du premier semestre est arrivée, l'intensité de l'effort et l'utilisation des caractéristiques individuelles qui pouvaient naître de la condition expérimentale ont été atténuées. Ainsi, le contrôle terminal n'a-t-il pas été mieux réussi par les étudiants du groupe traité.
Dans l'architecture du contrôle des connaissances, le contrôle terminal compte pour la moitié de la note finale d'une matière donnée. Il existe une part d'aléatoire dans les examens terminaux (Becker et Rosen, 1996) et cette période concentre tous les tests sur quelques jours avec des amplitudes horaires importantes. De surcroît, c'est la première expérience de partiels pour les étudiants. Ces derniers ne bénéficient pas d'entraînement au préalable, les formes d'évaluation pouvant sensiblement différer des formes d'examens en contrôle continu. Ces éléments peuvent provoquer un stress qui agirait négativement avec la capacité à investir davantage d'efforts en puisant dans des capacités spécifiques à l'enseignement supérieur.

Les étudiants auraient ainsi considéré que leurs notes de contrôle continu étaient suffisantes pour valider le semestre et ils ont donc réduit leurs efforts pour atteindre une note moins élevée lors du contrôle terminal. En d'autres termes, il suffisait pour eux d'atteindre une note de compensation suffisante pour valider la matière. C'est ce qui a dû se produire pour le contrôle terminal d'introduction à l'économie pour lequel la note moyenne s'élève à 7/20 pour le groupe traité. Ce dernier a atteint en moyenne la note de 13/20 pour le contrôle continu. Cela expliquerait pourquoi les étudiants sous la condition expérimentale n'ont pas eu un gain de

réussite lors du partiel d'introduction à l'économie. Il serait également possible d'expliquer l'écart de notes moyennes non significatif entre le groupe traité et le groupe témoin lors du contrôle terminal par le fait que l'examen en lui-même ne permettait pas d'appliquer des dispositions acquises au cours du processus expérimental.

Néanmoins, leurs résultats ont été meilleurs que le groupe témoin sur les autres matières pendant la période d'examens en contrôle terminal. Les étudiants du groupe traité ont ainsi mieux réparti leurs efforts d'apprentissage sur toutes les matières, afin de maximiser leurs chances de valider leur semestre. Cette répartition de l'effort peut être engendrée par une meilleure prise en considération de l'acquisition de connaissances pour chaque examen et une meilleure compréhension de ce qui était attendu d'eux.

Cette stratégie de réussite ne s'est cependant pas traduite par de meilleures chances de valider le premier semestre par rapport à des étudiants qui n'ont pas bénéficié du dispositif de partenariat évaluatif. Dans ce cadre, l'effort supérieur fourni lors des examens en contrôle continu n'a pas été suffisant pour compenser la baisse de l'effort aux examens terminaux. Les notes ont été meilleures, mais de façon générale la promotion a généralement validé son semestre, sans que ce dispositif apporte de chances supplémentaires.

2. La méfiance des étudiants de deuxième année

Les interprétations sont sensiblement différentes pour les étudiants de deuxième année. D'une part, ces étudiants connaissent déjà l'envers du décor de la formation ainsi que le jeu d'acteurs de l'équipe pédagogique et administrative. Leur comportement est par conséquent en adéquation avec cet environnement pour être en condition de réussite. Ils sont rompus aux règles qui entourent leur quotidien. En ce qui concerne la participation, ils ont ainsi internalisé les pratiques et le discours des enseignants sur la manière de la juger en travaux dirigés et les étudiants sont en mesure d'apprécier la validité et la portée de son évaluation. En d'autres termes, ils connaissent les problèmes d'efficacité et savent s'en accommoder dans leur cursus universitaire.

D'autre part, il faut considérer que pour la majorité des étudiants, leurs attentes et leur engagement envers le module de microéconomie sont faibles relativement à d'autres matières. Dans le discours des étudiants, c'est une discipline trop éloignée de la réalité et qui ne fournit pas des outils directement mobilisables dans un environnement professionnel. La conséquence de cette conception est un faible effort à l'apprentissage qui se traduit régulièrement par des résultats aux examens en dessous de la moyenne, en première année et en deuxième année. Ils abandonnent délibérément ce module pour se focaliser sur les autres.

Ainsi, lorsqu'un enseignant leur propose de participer à un partenariat dont l'objectif est de négocier la règle d'évaluation de la participation, ils se méfient de son discours et de sa véracité, autant qu'ils doutent de son impact sur leur réussite. Alors que dans les deux groupes expérimentaux de deuxième année, les étudiants ont accepté à l'unanimité de jouer le jeu de la négociation de l'évaluation de la participation, ils n'ont pas pris en considération les apports de cette démarche dans leur processus d'apprentissage.

La condition expérimentale a ainsi conduit les étudiants à davantage rester sur leurs acquis de connaissances antérieurs pour se conformer à la règle co-construite, sans penser qu'elle se traduirait par une véritable évaluation. Pour eux, il n'y avait ainsi pas assez d'incitation à fournir un effort pour ce module, puisque, quelle que soit la quantité d'efforts fournie, cela se traduirait par de l'échec aux examens.

De plus, la proposition de l'enseignant a pu être mal interprétée par les étudiants en contradiction avec leur socle de compréhension des règles de la formation. Cette nouveauté a pu engendrer un stress différent pour répondre à des attentes différentes à celles auxquelles ils étaient habitués à se conformer dans un contexte sans changement. Cette nouveauté a donc été mal perçue par des étudiants dans un module pour lequel ils estimaient que, quelles que soient les conditions du changement, elles ne modifieraient pas fondamentalement leurs manières d'apprécier son contenu et ses méthodes. Au contraire, cela les forçait à davantage s'impliquer dans un module où l'effort d'implication n'est de toute manière pas récompensé par un résultat positif.

Par conséquent, la condition expérimentale a détérioré la performance plutôt que de la maintenir ou de la renforcer lors des deux examens en contrôle continu.

En revanche, lorsque nous analysons le résultat au partiel, l'effet est inversé : les étudiants ont mieux réussi au contrôle terminal de microéconomie et ont également globalement mieux réussi aux examens terminaux de tous les autres modules. Ce renversement de situation peut s'expliquer par la prise en compte des étudiants de leur note de participation au module, c'est-à-dire la fin du processus expérimental. En effet, l'issue du partenariat se conclut par la note qui est calculée à partir des critères co-déterminés par le groupe d'étudiant et par l'enseignant. En moyenne, la note de participation obtenue par le groupe expérimental est de 13,88 ($\sigma^2 = 1{,}84$) ce qui correspond quasiment au double de la note obtenue lors des examens en contrôle continu. La remise de la note grâce à la distribution de la grille et des commentaires associés à l'évaluation de la participation a été vécue positivement par les étudiants qui ne s'attendaient pas à ce type de

résultat. Pour les étudiants qui ont mesuré l'ampleur de l'écart entre leurs résultats aux tests de connaissances et à cette évaluation de leurs autres aptitudes, cela a produit un changement de comportement important. En tenant son engagement, de respecter le partenariat, l'enseignant a ainsi conduit les étudiants à se remettre en question sur leurs stratégies de réussite.

Cette note aurait ainsi produit un changement de comportement dans l'effort d'acquisition de connaissances, aussi bien en microéconomie, pour tenter de rattraper les résultats du contrôle continu, mais également pour les autres examens. Cette expérience a permis de leur montrer qu'ils avaient les capacités de réussir et qu'il ne fallait pas demeurer dépendant au sentier de leur apprentissage, mais qu'ils devaient prendre en main leur parcours d'études. Leur effort d'apprentissage s'est soldé par une réussite plus importante que le groupe témoin aux examens terminaux.

In fine, cette remise en question a permis d'augmenter les chances de valider le troisième semestre pour les étudiants du groupe expérimental. Par conséquent, la note associée au partenariat évaluatif, la prise en compte des aptitudes acquises dans la formation pour réussir, ainsi que celles qui étaient offertes par le dispositif expérimental ont permis à ces étudiants de reprendre confiance en eux au cours de leur troisième semestre d'études, de remettre en cause le processus d'échec qui semblait inévitable et cela s'est traduit par davantage de réussite.

3. Une interprétation générale

Si nous devions synthétiser nos résultats en une formule simple, celle-ci serait que les effets de la mise en place d'une négociation des règles d'évaluation entre les étudiants et les enseignants pour promouvoir la réussite dépendent du contexte dans lequel elle se réalise.

Si l'on se réfère aux arguments de la théorie économique, le partenariat et la coopération peuvent aboutir à une situation socialement optimale si les deux parties de la relation considèrent que l'enseignement est un bien public, c'est-à-dire que les deux parties contribuent à sa production et qu'elle n'entraîne pas de rivalité dans sa consommation. Nous pouvons légitimement tracer le parallèle de cette situation avec la relation étudiant-enseignant en première année de licence. À ce niveau, les étudiants ont dans leur majorité « choisi » de venir dans ce type de cursus d'enseignement supérieur pour plusieurs raisons : l'accueil sans sélection de tous les profils de bacheliers, le temps effectif en présentiel réduit, la possibilité de réfléchir pour se réorienter, la localisation géographique, etc. Bien que leurs attentes individuelles soient différentes, elles sont néanmoins tournées vers l'acquisition de connaissances propres à la formation et à la discipline pour avancer dans leur formation supérieure et vers leur choix de carrière professionnelle. De

plus, la peur de ne pas réussir renforce leur motivation à éviter cette situation. De surcroît, le module d'introduction à l'économie est un module fondamental de leur licence dont les connaissances transmises permettront d'appréhender le reste de la formation en économie. Les étudiants sont donc en attente de l'instruction que leur offrira l'équipe enseignante, laquelle conçoit par définition un intérêt à délivrer cette instruction. En considérant l'évaluation de la participation, les étudiants n'ont pas de croyances négatives sur la façon dont elle est évaluée. Dans ce contexte, l'opacité liée à la nouveauté quant aux règles de fonctionnement permet de leur faire accepter des dispositifs qui leur semblent opportuns pour leur stratégie de réussite à l'université.

Par conséquent, le pouvoir de négociation a été accepté et le processus de négociation a été validé par les étudiants. L'exécution du contrat a été crédible et cela a engendré un effort d'apprentissage constant au cours des examens en contrôle continu. La vérifiabilité du contrat a été permise par la note de participation à la fin du processus expérimental. L'optimum social peut être donc être atteint dans le contexte d'initialisation de la formation lorsque les étudiants et les enseignants entrent en interaction pour dialoguer et contribuer à une règle commune d'évaluation.

En revanche, les étudiants qui sont arrivés en deuxième année ont franchi des étapes qui ont manifestement modifié leurs attentes sur la formation. Ils ont ainsi davantage compris comment fonctionnait l'université, dans laquelle se réalise le jeu d'acteur de leurs enseignants. Ils ont compris que l'évaluation de la participation ne se réalise pas de manière efficace. De plus, en ce qui concerne le module de microéconomie, les étudiants ne le considèrent pas comme faisant partie intégrante de leur réussite dans la formation, tandis que les enseignants le considèrent comme étant une matière fondamentale de la discipline. Par conséquent, la relation entre les étudiants et les enseignants se transforme. Cette situation est alors plus favorable à une divergence d'intérêts quant à la formation, ce qui correspond à la deuxième manière d'appréhender d'un point de vue théorique la relation enseignant-étudiant.

Dans ce contexte, bien que l'expérimentation ait donné l'opportunité de négocier un contrat d'évaluation de la participation et que les étudiants aient accepté d'y participer, ces derniers n'ont pas admis que l'exécution du contrat serait crédible. La négociation n'a alors pas engendré de réduction de la croyance négative sur les règles d'évaluation et n'a pas permis de donner davantage d'intérêt à l'effort d'apprentissage dans ce module. De plus, le fait que l'exécution du contrat n'ait pas été rendue visible avant la fin du semestre a pu conforter les étudiants dans leurs croyances.

À la différence des étudiants de première année, c'est donc davantage la manière dont l'issue de la négociation s'est réalisée qui a fait changer les comportements sur l'effort d'apprentissage et sur la capacité de réussir aux examens que le fait de se saisir de cette négociation pour adopter des dispositions favorables à la réussite. Lorsque le contrat a été exécuté par l'enseignant, c'est-à-dire qu'il a tenu l'engagement pris avec les étudiants de tenir compte de leurs critères d'évaluation pour évaluer leur participation, les étudiants ont alors pris conscience de la véracité du contrat et se sont remis en question. Finalement, cette remise en question a permis aux étudiants de renforcer leur motivation et de croire en leurs capacités y compris dans un module sur lequel ils ne portaient pas d'intérêt pour leur réussite et ils ont donc fourni un effort plus important lors du contrôle terminal de connaissances.

Ainsi, l'expérimentation permet de compléter les apports de la littérature sur la négociation dans un cadre d'aléa moral en fournissant l'idée selon laquelle l'issue de la négociation, le contrat, doit être visible par les parties prenantes pour qu'il puisse rendre crédible l'accord et réduire les croyances négatives sur l'autre partie. Nous rentrons ainsi dans une dimension dynamique de la négociation des contrats et de la vérification de l'effort en présence de divergences d'intérêts.

Enfin, nous avons explicité que le contexte des interventions en science de l'éducation était généralement favorable à leurs implémentations. Nous montrons ainsi que lorsqu'il n'est pas favorable, cela peut conduire à des effets non anticipés et qu'il peut détériorer la réussite. Par conséquent, la négociation des règles d'évaluation produirait un meilleur effet en début de processus de formation lorsque les attentes des deux parties de la relation sont sur le même niveau, c'est-à-dire la contribution à un bien commun, parce qu'elle permet de comprendre les règles en vigueur et de promouvoir des environnements plus propices à la parole des étudiants pour améliorer ce que leurs enseignants proposent d'évaluer. Lorsque la négociation est implantée au cours du processus de formation, une condition d'efficacité serait de rendre visible l'exécution du contrat pour réduire les croyances négatives.

VIII. Conclusion

Nous avons conduit un partenariat expérimental d'évaluation dans le cadre de deux modules de formation de la licence d'Économie-Gestion de l'Université Paris-Est Marne-la-Vallée. Ce partenariat d'évaluation consistait à négocier les critères d'appréciation de la participation des étudiants dans les travaux dirigés. Nous avons mesuré l'impact de ce processus sur les notes obtenues dans le module et sur les résultats dans les autres composantes de la formation. Nous

montrons que la négociation produit un effet positif sur la réussite quand la relation entre l'enseignant et les étudiants est favorable à la négociation. L'effet positif sur la réussite dépend également de la visibilité de l'exécution du contrat dans le cas où la relation entre les enseignants et les étudiants est constituée de divergences d'intérêt.

1. Des apports à l'économie des organisations

Nous contribuons ainsi à une démarche générale d'expérimentations de dispositifs organisationnels innovants qui permet de tester les conditions de leur mise en œuvre à une échelle globale et leurs impacts à divers degrés (Shaw, 2009). Ces expérimentations permettent de se prémunir de plusieurs biais de mesure engendrés par exemple par la sélection non aléatoire d'échantillons d'individus qui ont des caractéristiques qui interfèrent avec les effets des dispositifs que l'expérimentateur souhaiterait identifier. Nous renforçons ainsi l'idée qu'il est possible de conduire ces expérimentations dans des conditions réelles, en dehors d'un laboratoire, en l'occurrence au sein de l'université, tout en se prémunissant des limites d'ordre éthique. La réussite d'un tel processus expérimental tient notamment à l'implication des parties prenantes de l'organisation à plusieurs étapes du protocole.

D'une part, nous apportons une contribution à l'ensemble des études de mesures d'impact de dispositifs de remédiation dans l'enseignement supérieur (Sneyers et De Witte, 2017). Nous complétons ainsi le panorama des outils à la disposition des responsables de formation et des enseignants pour améliorer la conduite des étudiants dans leur parcours de formation, à l'instar du tutorat, des bourses de mobilité ou sur critères sociaux, des enseignements de méthodologie, des cours de renforcement, des règles de compensation d'unité d'enseignement, etc. Nous développons l'idée selon laquelle l'accompagnement des étudiants, ayant des profils et un passé éducatif de plus en plus hétérogènes, doit s'outiller de plusieurs techniques que les acteurs de la formation peuvent utiliser.

Dans ce cadre, nous renforçons l'idée que l'un des déterminants de l'engagement des étudiants dans leur apprentissage se noue dans la relation qu'ils créent avec l'équipe enseignante (Correa et Gruver, 1987). Ces interactions doivent s'incarner au travers de dialogues et de coopérations entre les deux pôles de la relation, afin de promouvoir la parole et les décisions des étudiants qui sont toutes autant valables que celles des enseignants (Jarvis, Dickerson et Stockwell, 2013). Par conséquent, le poids attribué à l'enseignant dans la formation peut être restreint afin de favoriser l'émergence de dispositions nouvelles de la part des étudiants, dispositions que nous avons citées à plusieurs reprises comme étant constitutives de l'apport de la formation supérieure : l'autonomie, la prise de décision, la réflexion critique.

D'autre part, nous montrons qu'il est possible d'expérimenter des outils innovants pour améliorer les pratiques d'évaluation qui soutiennent l'effort d'apprentissage davantage qu'elles ne sanctionnent sa progression (McMillan, 2015). L'inertie de l'évaluation à l'œuvre à l'université n'est ainsi pas immuable, elle peut au contraire être renversée pour incorporer plus de considérations « formatives ». Il existe des marges de progression et les organisations d'enseignement supérieur peuvent s'appuyer sur le partenariat évaluatif pour renforcer la fiabilité et la validité des examens. Dans ce contexte, l'évaluation peut être transformée et dépendante de préconisations formulées tant par les étudiants et que par les enseignants. Une meilleure technique d'évaluation et de contrôle des connaissances et des compétences permet d'obtenir une meilleure qualité du signal émis sur la production des diplômes.
Le message porté par le partenariat évaluatif est que la conception des tests doit aussi faire partie intégrante de la conception de l'enseignement.

Cette expérimentation répond en partie à des préoccupations générales sur le modèle évaluatif de l'université. En reprenant les termes de l'Inspection générale de l'administration de l'éducation nationale et de la recherche (Gauthier et al., 2007), il s'agit de promouvoir une matrice évaluative, davantage plastique en fonction des publics accueillis et des acquis visés, qui permette de construire l'évaluation concomitamment avec l'enseignement et qui soit tournée vers une évaluation des compétences en complément des connaissances.

D'un point de vue de microéconomie des organisations, nous apportons des arguments favorables à la mise en place d'environnements de négociation des dispositifs de vérification de l'information. Cette négociation procure des marges de manœuvre qui incitent à l'effort pour atténuer l'aléa moral sur les positions individuelles. Il est possible de donner du pouvoir à l'agent sans qu'il opte pour un comportement opportuniste et de réduire le pouvoir du principal afin qu'il n'abuse pas de sa position dominante. Cet environnement de négociation peut aboutir à des contrats qui sont constitutifs d'une allocation optimale des ressources entre les parties de la relation.

2. La portée des résultats est limitée par le protocole expérimental

La recherche que nous avons conduite n'est cependant pas exempte d'un certain nombre de limites qui peut réduire la pertinence de son message.
Premièrement, les caractéristiques de l'expérimentation soulèvent un problème de validité externe. La taille des effectifs, qui est faible en comparaison d'études de plus grande envergure (par exemple Dmitrijeva et al., 2014), engendre nécessairement des précautions quant à la volonté

de généraliser à l'ensemble de la population étudiante les effets que nous mesurons. Le fait qu'elle ait été réalisée au sein d'une unique unité de formation et de recherche, dans une seule université, restreint également la portée des résultats que nous pourrions obtenir dans un environnement de formation différent qui comporte des modalités de fonctionnement distinctes, voire opposées. Nous rappelons également que les partenariats évaluatifs ont été menés avec un seul enseignant et que par conséquent la réplicabilité de ces processus demeure conditionnée par les pratiques et les croyances individuelles sur les normes d'évaluation à adopter, c'est-à-dire que dans ce contexte il serait également important de mettre en lumière les dispositions des enseignants à réaliser la négociation. Enfin, la temporalité du dispositif pose une question sur la durabilité des effets constatés.

Deuxièmement, la condition expérimentale est un processus défini comme la suite continue d'opérations aboutissant à un résultat davantage que l'intervention d'un procédé unique. Par conséquent, il existe des impacts sous-jacents engendrés par la condition expérimentale que nous ne sommes pas en capacité d'appréhender. Ainsi, en l'absence d'une modélisation plus fine de ce processus, nous avons délibérément évacué des dimensions qui renforcent ou atténuent l'impact sur les indicateurs de réussite. Par exemple, nous n'avons pas tenu compte du degré d'autonomie développé par les étudiants grâce à ce dispositif. Nous avons considéré que cette disposition s'était davantage produite chez le groupe expérimental et qu'elle permettait de générer un plus grand effort d'apprentissage et une réussite plus importante. La complexité du traitement et ses interactions avec d'autres variables sont par conséquent masquées.

Cette remarque nous permet également de parvenir à l'idée selon laquelle le phénomène de l'échec/de la réussite est protéiforme et que toutes ses dimensions n'ont pas été investiguées. Il reste un enjeu important de formalisation de la réussite qu'il ne faut pas réduire à deux indicateurs. Il serait donc nécessaire d'identifier d'autres variables de résultats pour apprécier les effets de la négociation des règles de l'évaluation dans leur ensemble.

Dans ce cadre, cette étude souffre de ne pas proposer de comparaison de l'issue des partenariats, à savoir le « contrat » contenant les règles d'évaluation et les notes obtenues par les étudiants, avec ce qui se pratique dans les autres modules de formation. Nous n'en avons pas tenu compte, car il n'était pas possible de comparer les niveaux de participation obtenus par les étudiants des autres groupes de travaux dirigés en raison d'échelles de mesure différentes et de critères de jugements variés. Par conséquent, la confrontation des résultats de participation dans les deux groupes d'étudiants n'aurait pas permis de produire des interprétations quant au niveau effectif de participation.

Enfin, une limite de cette recherche tient à la nature des partenariats offerts aux étudiants. Deeley (2014) a en effet déterminé que les coopérations peuvent être fortes, lorsque les étudiants et les enseignants conviennent d'une règle d'évaluation et/ou réalisent conjointement l'évaluation, ou faibles, lorsque les critères d'examen ne sont pas coconstruits et/ou que l'évaluation n'est pas conduite par les deux parties. Dans notre protocole expérimental, le partenariat correspond davantage au modèle « faible », dans le sens où les étudiants n'ont pas utilisé la grille pour s'auto-évaluer sur la participation, l'évaluation restant sous la responsabilité du chargé de travaux dirigés. L'engagement des étudiants n'allait donc pas au-delà de la participation à la discussion et à la formalisation des critères de notation. De plus, la négociation n'a pas concerné les examens de contrôle des connaissances, pour lesquels les enjeux sont plus importants : ils structurent davantage le cursus de la formation parce qu'ils font partie intégrante du rituel de l'apprentissage et récompensent davantage l'effort des étudiants. Par conséquent, la pertinence de la négociation portant sur la participation pourrait être remise en question.

3. Cette expérimentation est une première phase d'exploration pour de futures pistes de recherche

L'ensemble de ces remarques permet de dresser de nouveaux jalons qui offriraient de perspectives inédites en vue de futures recherches. Dans la mesure où nous pouvons considérer cette expérimentation comme une invitation à d'autres initiatives expérimentales, il serait judicieux de changer d'échelle afin d'incorporer un plus grand nombre de parties prenantes et produire de nouveaux résultats.

Il est souhaitable d'approfondir l'idée de la négociation de l'évaluation dans un cadre plus ambitieux, par exemple au niveau d'un examen de connaissances en contrôle continu, voire en contrôle terminal. Tout l'enjeu réside alors dans la délimitation de la marge de négociation des parties prenantes pour ne pas amoindrir le caractère certificatif de l'évaluation des connaissances des étudiants et ne pas engendrer des comportements que les responsables de formation voudraient éviter. L'objectif serait de continuer à développer les aptitudes et les compétences issues de la formation supérieure en symbiose avec l'objectif d'acquisition de connaissances disciplinaires.

Une voie ouverte de recherche serait de quantifier d'autres variables de résultats à l'issue de ces dispositifs. Ces variables seraient en phase avec l'idée du développement d'aptitudes propres à l'enseignement supérieur et pourraient ainsi s'inscrire durablement dans une économie de la

connaissance qui mobilise les facultés d'apprentissage et d'évaluation individuelle tout au long de la vie.

Dans le but d'approfondir la notion de négociation de la règle d'évaluation, il serait judicieux de déterminer dans quelle mesure les différentes formes de contrats qui découlent de ces négociations influencent à la hausse ou à la baisse les indicateurs d'intérêt. Ainsi, il serait possible de proposer des choix d'évaluation, sous la forme de menu, dans lesquels les étudiants se positionneraient. À l'instar du choix du module, les étudiants choisiraient leurs évaluations.

Enfin, au moment où la loi relative à l'orientation et la réussite des étudiants dans les universités instaure l'obligation de signature d'un contrat pédagogique liant l'étudiant à son cursus de formation, dans quelle mesure l'incorporation au sein de ces contrats de stipulations liées à l'évaluation renforcerait-elle leur crédibilité et permettrait-elle de faire adopter des dispositions individuelles nouvelles ?

Annexes du chapitre 3

4. Réussite à l'université en France

Nous avons compilé plusieurs rapports d'information du Ministère de l'Enseignement supérieur pour dresser l'évolution du taux de diplomation en trois années en licence, depuis la généralisation du diplôme en 2004. Les dernières statistiques descriptives disponibles correspondent à la promotion d'étudiants de l'année 2012-2013 pour laquelle il est possible de déterminer son devenir trois années ou quatre années après la première inscription. Les indicateurs les plus révélateurs du phénomène de la réussite sont inscrits dans le tableau A1.
Ces indicateurs n'offrent qu'une vision partielle de la réussite à l'université parce que d'une part, ils ne tiennent compte que des seules réorientations au sein de l'université et d'autre part, ils sont calculés sur la base des inscriptions administratives. Par conséquent, les réorientations vers des filières autres qu'universitaires (paramédicale ou sociale, école d'ingénieurs…) et la présence effective des étudiants en cours de formation ne sont pas utilisées.
De plus, chaque université possède sa propre population étudiante, c'est-à-dire que les caractéristiques sociodémographiques et les parcours scolaires antérieurs des étudiants diffèrent selon les établissements. Or, la probabilité de réussite (ou de passage) apparaît corrélée avec des caractéristiques telles que l'origine sociale et le parcours scolaire antérieur (série du baccalauréat, âge au moment de son obtention, ancienneté de son obtention, la mention obtenue). Les caractéristiques non observables de chaque établissement, par exemple les modalités d'évaluation et les dispositifs d'accompagnement des étudiants sont également des facteurs qui influencent la réussite.

Tableau A1 : Indicateurs de réussite d'obtention de la licence en trois ans

Promotion	2004-2005	2005-2006	2006-2007	2007-2008	2008-2009	2009-2010	2010-2011	2011-2012	2012-2013
Taux de licenciés en 3 ans	28,8*	27,9	27,8	27	27,2	26,5	27,8	27,2	28,4
Part de femmes	62	n.c.	n.c.	62,2	61,6	60,3	59,7	59,3	58,9
Part de bacheliers Général	73,1	n.c.	n.c.	75,3	74,2	71,4	71	69,2	75,3
Part de bacheliers Technologique	16,9	n.c.	n.c.	16,1	n.c.	n.c.	n.c.	n.c.	n.c.
Part de bacheliers Professionnel	4	n.c.	n.c.	4,2	n.c.	n.c.	n.c.	n.c.	n.c.
Taux de femmes licenciées en 3 ans	33	n.c.	n.c.	30,6	30,6	30,1	31,7	31,1	32,9
Taux d'hommes licenciés en 3 ans	21,9	n.c.	n.c.	21,1	21,7	21,2	21,9	21,6	21,8
Taux de bacheliers Général licenciés en 3 ans	36	n.c.	n.c.	33	33,3	33,2	34,7	35	35,8
Taux de bacheliers Technologique licenciés en 3 ans	7,4	n.c.	n.c.	9,8	7,5	8,7	9,8	9,2	8,1
Taux de bacheliers Professionnel licenciés en 3 ans	2,5	n.c.	n.c.	2,7	3,1	3	3,7	3,1	2,3
Redoublement de la première année	24,6	n.c.	n.c.	24,1	25	n.c.	n.c.	n.c.	n.c.
Non-réinscription après la première année	27,8	n.c.	n.c.	31,8	30	n.c.	n.c.	n.c.	n.c.
Nombre de néo-entrants à l'université	179300	172200	166100	154300	150100	n.c.	n.c.	n.c.	n.c.

Note : * parmi les 179300 néo-entrants à l'université lors de l'année 2004-2005, 28,8 d'entre eux ont obtenu une licence trois années plus tard. n.c : non communiqué.

Sources :

Fouquet S., (2013), Parcours et réussite en licence et en master à l'université, Note d'information n°13-02, MESR-DGESIP-DGRI, Service d'information et études statistiques (SIES).

Fouquet S., (2014), Parcours et réussite aux diplômes universitaires : les indicateurs de la session 2012, Note Flash n°08-04, DGRI-SCSER, Service d'information et études statistiques (SIES).

Maetz I., (2015), Parcours et réussite aux diplômes universitaires : les indicateurs de la session 2013, Note Flash n°07-05, MENESR DGESIP/DGRI — SCSESR - Service d'information et études statistiques (SIES).

Maetz I., (2016), Parcours et réussite aux diplômes universitaires : les indicateurs de la session 2014, Note Flash n°02-01, MENESR DGESIP/DGRI — SCSESR - Service d'information et études statistiques (SIES).

Maetz I., (2016), Parcours et réussite aux diplômes universitaires : les indicateurs de la session 2015, Note Flash n°11-15, MENESR DGESIP/DGRI — SCSESR - Service d'information et études statistiques (SIES).

Maetz I., (2017), Parcours et réussite en licence et en Paces : les résultats de la session 2016, Note Flash n°11-18, MESRI — Service d'information et études statistiques (SIES).

Prouteau D., (2009), Parcours et réussite en licence des inscrits en L1 en 2004, Note d'information n°09-23, MESR-DGESIP-DGRI, Service d'information et études statistiques (SIES).

5. Grille d'évaluation de la participation à l'issue de la négociation

Nous présentons à titre d'illustration du protocole expérimental, un exemple de grille d'évaluation qui a été formulée à l'issue des trois séances de réflexion collective sur les critères de notation de la participation. Ces grilles ont été communiquées aux étudiants à la quatrième séance à l'issue de la validation pour qu'ils en conservent un exemplaire tout au long du semestre. L'évaluation finale de la participation a été transmise aux étudiants en entourant chaque item d'évaluation et en communiquant la note finale qui tenait compte des critères. Pour rappel, chaque item avait la même pondération.

Grille d'évaluation de la note de participation en Travaux dirigés Introduction à l'Economie

GROUPE de TD n° XXX

Nom : XXX

Règles		Fréquence			Commentaire
Étudiants	Poser des questions (y compris si l'on pense que ce n'est pas cohérent) - Sur les mécanismes - Interprétations Vocabulaire	JAMAIS	PARFOIS	RÉGULIÈREMENT	
	Réponse aux questions même si on pense avoir faux	JAMAIS	PARFOIS	RÉGULIÈREMENT	
	Être présent	JAMAIS	PARFOIS	RÉGULIÈREMENT	
	Être ponctuel	JAMAIS	PARFOIS	RÉGULIÈREMENT	
	Être attentif.ve	JAMAIS	PARFOIS	RÉGULIÈREMENT	
	Être polis	JAMAIS	PARFOIS	RÉGULIÈREMENT	
Chargé de TD	Apporter son matériel de travail (résumé, texte, ordinateur, notes de cours …)	JAMAIS	PARFOIS	RÉGULIÈREMENT	
	Travailler pendant le TD (prise de notes – prise de la correction)	JAMAIS	PARFOIS	RÉGULIÈREMENT	
	Prises d'initiatives pendant le TD (bien se sentir en TD)	JAMAIS	PARFOIS	RÉGULIÈREMENT	
	Participer au débat / Réagir aux interventions de ses camarades	JAMAIS	PARFOIS	RÉGULIÈREMENT	
	Interaction avec le chargé de TD pour approfondir	JAMAIS	PARFOIS	RÉGULIÈREMENT	

Note :

6. Test d'équivalence des notes de participation

Afin de s'assurer que la condition expérimentale a été homogène entre les groupes expérimentaux, nous recourrons à une ANOVA. Ce test permet de détecter si une note moyenne est significativement différente des autres. Nous utilisons la statistique de Fisher. Le test ne permet pas de rejeter l'hypothèse de nullité des écarts entre toutes les notes moyennes. Par conséquent, il n'y a pas eu de surnotation ou de sous-notation de la participation à l'issue de la négociation de la règle d'évaluation.

Tableau A2 : Test d'homogénéité des notes moyennes de participation des groupes traités

Groupe	Nombre d'observations	Note de participation moyenne	Test de Fisher
Groupe traité n°1 L1	29	12,972 (2,028)	1,86 p-value = 0,1204
Groupe traité n°2 L1	31	12,710 (2,132	
Groupe traité n°3 L1	26	13,077 (3,068)	
Groupe traité n°4 L2	32	13,945 (2,225)	
Groupe traité n°5 L2	29	13,822 (1,382)	

Note : les erreurs-types sont entre parenthèses. Le test de Fisher teste l'hypothèse nulle de l'existence d'une différence entre tous les moyennes des notes de participation. Les moyennes sont calculées sur tous les étudiants ayant été dans la condition expérimentale.

Conclusion générale

L'économie des contrats complets prédit qu'au sein d'une relation d'agence, en présence d'aléa moral, un dispositif de vérification de l'information permet d'aligner les intérêts des deux contractants. En effet, en raison de sa capacité à fournir des renseignements sur l'activité de l'agent, le dispositif de vérification de l'information est un support de la mise en œuvre des incitations, qu'elles soient sous la forme de récompenses ou de sanctions. Dans les organisations productives, l'existence de ce dispositif répond à la fois à un problème de coopération entre les individus et à un problème de coordination des activités.

Cependant, l'utilisation du dispositif de vérification de l'information peut produire des effets ambivalents. Au lieu de discipliner des comportements opportunistes, ce dispositif peut engendrer des coûts cachés et réduire la motivation intrinsèque des agents à réaliser une activité qui leur a été attribuée. En définitive, sous certaines conditions, il génère une réduction du résultat final, et une perte en termes d'allocation ce qui nuit à l'efficacité.

Ces intuitions ont été formulées dans des cadres théoriques diversifiés et elles ont également été confirmées par des preuves empiriques, depuis les années 1970. Néanmoins, l'économie des dispositifs de vérification de l'information est restée bornée à une relation d'agence traditionnelle : la relation d'emploi. Pourtant, des asymétries informationnelles et des divergences d'intérêts sont présentes dans la plupart des relations contractuelles, qu'elles soient explicites et implicites.

Il nous a donc semblé opportun d'étendre le spectre d'analyse des dispositifs de vérification de l'information afin, d'une part, de vérifier leur pertinence pour des relations d'agence alternatives à la relation d'emploi, d'autre part, de mesurer leurs effets bénéfiques ou leurs effets préjudiciables sur la répartition des ressources, ainsi que de distinguer les éléments qui composent leurs réglages.

En définitive, il s'agissait d'ouvrir la « boîte noire » des dispositifs de vérification de l'information au sein des organisations et de déterminer l'existence d'effets éventuellement ambivalents sur le comportement des parties de la relation d'agence.

Dans cette thèse de doctorat, nous nous sommes appuyés sur des expérimentations contrôlées de terrain avec assignation aléatoire de plusieurs variations du dispositif de vérification de l'information. Cette méthode permet de produire une analyse plus robuste des liens de causalité entre un changement d'environnement et des réactions comportementales. Le recours à ces protocoles de recherche nous a permis d'étudier trois relations d'agence particulières dans trois terrains différents : la relation d'emploi dans un centre d'appels sous-traitant ; la relation de contribution sociale dans l'agence de recouvrement des cotisations sociales ; la relation d'enseignement au sein d'une université.

À chaque relation correspond un dispositif de vérification de l'information et un problème économique sous-jacent, c'est-à-dire l'existence d'un aléa moral. Ainsi, l'objectif de cette thèse de doctorat était de pouvoir mesurer l'impact d'une modification du dispositif de vérification de l'information afin de résoudre un problème organisationnel. Les résultats que nous avons obtenus ont été, la plupart du temps, inattendus.

Afin de conclure notre propos, nous présentons les principaux résultats de ces expérimentations de terrain et nous ouvrons des perspectives de recherche pour nous permettre de dépasser les écueils qui se sont manifestés au cours de ce travail de microéconomie appliquée.

I. Les principaux résultats de cette recherche

Dans le premier chapitre, à partir des méthodes appliquées de l'économie publique, nous avons expérimenté l'explicitation du dispositif de contrôle de la fraude au prélèvement social en partenariat avec l'Acoss, afin d'augmenter le consentement à la contribution sociale des entreprises. Cette expérimentation testait les effets d'un envoi de message ciblé qui expliquait les conditions générales de la vérification de la déclaration sociale, dont le contenu était modifié pour contenir des renseignements spécifiques sur l'existence de sanctions ou pour insister sur la moralité contributive. Après l'envoi simultané de dix messages différents à des groupes de PME (effectif inférieur à 250 salariés), nous avons mesuré les variations pour des variables de résultats en stock et en flux issues des déclarations sociales des entreprises.
Nous avons constaté une réduction du nombre de créations d'emplois en CDD à l'issue de l'envoi de trois messages : l'association de la fraude au vol ; la description des sanctions non financières ; le respect de la norme sociale du prélèvement obligatoire.
Ainsi pour que les entreprises consentent davantage au prélèvement social, le message qui explique les conditions du contrôle gagne en efficacité lorsqu'il est associé à des considérations morales. Par conséquent, il semble que pour améliorer le consentement, les entreprises doivent comprendre l'utilité et la légitimité d'un tel dispositif.

Dans le deuxième chapitre, à partir des conclusions de la littérature en économie comportementale, en psychologie sociale du travail et en sociologie du travail, nous avons expérimenté la réduction de l'intensité de la surveillance électronique dans un centre d'appels sous-traitant, afin d'améliorer la qualité de vie au travail des conseillers téléphoniques. Cette expérimentation testait les effets de la réduction du nombre d'informations collectées sur la performance des conseillers et la limitation des informations disponibles en temps réel pour les superviseurs du plateau téléphonique, ainsi que des alertes automatiques de dépassement de

temps de communication. Nos variables de résultats étaient à la fois l'effort objectif et la perception de la qualité des conditions de travail des salariés. Pendant seize semaines, le traitement expérimental alternait entre un groupe test et un groupe témoin qui changeaient de composition tous les quinze jours.

À l'aide d'un protocole mixte de recherche, quantitatif et qualitatif, nous avons mesuré une augmentation significative de la fatigue et du stress ressenti, ainsi qu'une perception d'une surveillance accrue pour les salariés ayant bénéficié du traitement. Cet effet contre-intuitif a été interprété à l'aune de deux considérations.

D'une part, les conseillers téléphoniques ont été déstabilisés par la suppression de l'accès à des indicateurs qui leur permettaient de se situer dans leur travail quotidien. Le dispositif de vérification de l'information, en créant des indicateurs individuels, permet de formaliser la production dans un secteur où l'activité est par nature immatérielle. Il facilite la gestion du temps, qui semble être une fonction perçue comme primordiale pour rendre un service de qualité.

D'autre part, les superviseurs ont semblé substituer davantage de surveillance physique à leur temps de surveillance électronique en étant davantage présents sur le plateau téléphonique afin de collecter les renseignements qui n'étaient plus à leur disposition. Plutôt que de permettre une meilleure coordination de l'activité, les superviseurs ont été jugés plus coercitifs dans leurs pratiques.

Dans le troisième chapitre, à partir de la microéconomie de l'enseignement et d'un courant émergent des sciences de l'éducation, nous avons expérimenté la négociation de la règle d'évaluation de la participation dans plusieurs modules de formation de l'UFR d'Économie-Gestion de l'Université Paris-Est Marne-la-Vallée. Cette expérimentation testait les effets d'une négociation entre l'enseignant et les étudiants des critères d'évaluation de la participation en travaux dirigés sur leur réussite aux épreuves semestrielles. Au début du semestre, les étudiants définissaient une liste de critères d'examen qui était amendée et validée collectivement.

Les étudiants de première année ont ainsi eu un gain significatif de points au cours du contrôle continu, mais pas lors du contrôle terminal et l'expérimentation n'a pas engendré de chances supplémentaires pour valider leur semestre. Tandis que pour les étudiants de deuxième année, leurs résultats de contrôle continu ont été réduits, mais leurs chances d'admission ont été améliorées.

Ces effets hétérogènes s'expliquent notamment par le contexte qui environne chaque relation d'enseignement. Pour les uns, la nouveauté de l'organisation a favorisé l'adoption de la négociation et a permis de se conformer aux attentes de l'enseignant, au niveau du contrôle continu. Pour les autres, la non-visibilité de l'exécution du contrat négocié a réduit la motivation à se conformer aux attentes de l'enseignant.

II. Les principaux apports à l'issue de cette recherche

Bien que ces trois recherches aient été conduites dans des organisations différentes et pour répondre à des questions distinctes, il est néanmoins possible de dresser des points de convergence qui sont autant de contributions à l'économie des dispositifs de vérification de l'information.

Premièrement, nous avons identifié trois types de dispositifs de vérification de l'information dans trois domaines d'activité qui correspondent à trois relations d'agence. Dans la relation d'emploi d'un centre d'appels, il s'agit d'un outil de surveillance électronique de l'effort au travail qui mesure en temps réel le temps passé au téléphone par les conseillers. Dans la relation de prélèvement des cotisations sociales, il s'agit d'une politique de contrôle comptable sur pièces ou de visites d'inspecteurs chargés d'enquêter sur la véracité des informations sociales fournies par les entreprises. Dans la relation d'enseignement, il s'agit de l'examen des connaissances ou des compétences qui est consacré à évaluer le degré d'acquisition des étudiants en vue de certifier leur niveau d'apprentissage.

Pour chacune de ces relations, nous avons donc pu montrer dans quelle mesure le dispositif de vérification de l'information était multidimensionnel et de quelle manière il était configuré. Son réglage dépend à la fois du niveau d'information collectée, qui définit l'intensité de la vérification, de la manière dont est explicité son fonctionnement et la justification de son implémentation, ainsi que de la manière dont il est accepté par les contractants de la relation. Ces trois dimensions ne sont pas les seules à caractériser dans la relation d'agence, mais nous avons souligné leur importance dans nos trois terrains de recherche.

La spécificité de ces dispositifs est qu'ils révèlent une facette de la relation d'agence : le pouvoir de la vérification est aux mains du principal. Par conséquent, dans une relation caractérisée par un aléa moral sur le comportement de l'agent, c'est-à-dire une asymétrie informationnelle en défaveur du principal, l'activation d'un dispositif de vérification revient à générer une asymétrie de pouvoir de contrôle de l'information en défaveur de l'agent. *In fine*, ce pouvoir peut être utilisé de façon opportuniste par le principal ce qui entraîne des effets ambivalents dans le résultat produit par l'agent : un comportement d'évitement de la tâche assignée, une baisse de la motivation intrinsèque, une culture du résultat plutôt qu'une aide à l'apprentissage. La vérification de l'information peut entraîner une forme de coercition des comportements en vue d'une normalisation de l'activité plutôt qu'une forme de détection des erreurs en vue d'une amélioration de l'activité. Cette position du principal est d'autant plus affirmée que pour chacune de ces

relations, une forme de hiérarchie est discernable et que cette dernière est renforcée par la vérification de l'information.

Après avoir mis à jour ce rapport de force, dans un deuxième temps, nous avons modifié le paramétrage des dispositifs de vérification de l'information. Ces modifications devaient permettre de résoudre un problème économique spécifique à chaque organisation, problème engendré par une potentielle interférence entre le dispositif de vérification de l'information et l'activité des agents. Ces changements micro-organisationnels permettaient de remettre en question la place du dispositif de vérification tout en laissant inchangé le système d'incitations : il était donc question de reformuler le contrat d'agence. Les réactions à ces changements ont été mesurées du point de vue des agents qui étaient susceptibles d'obtenir de plus grandes marges de manœuvre et pour lesquels le dispositif de vérification de l'information influençait les décisions. En définitive, ces variations étaient de nature à entraîner des bénéfices pour les deux parties de la relation, tout en conciliant les deux versants du dispositif de vérification de l'information : l'aspect disciplinant et l'aspect bienveillant.

D'une certaine manière, ces modifications destinées à résoudre un problème organisationnel ont permis de rendre visible le dispositif de vérification et ont donné des marges de manœuvre et de nouvelles possibilités d'arbitrage pour les agents. Ces derniers ont pris conscience de l'importance d'un tel outil dans la relation d'agence, que ce soit pour eux-mêmes ou pour le principal. Les employés utilisent la surveillance de leur temps de travail pour planifier leur activité, les cotisants s'appuient sur les règles de contrôle de la fraude sociale pour la gestion de leur main-d'œuvre temporaire et les étudiants s'inspirent des règles d'évaluation pour se positionner dans leurs apprentissages. Cette visibilité devait leur faire prendre conscience des potentielles décisions néfastes prises à leur propre initiative afin d'entreprendre des actions opposées tout en ayant davantage de latitude décisionnelle pour réaliser leur choix. Réduire l'intensité de la surveillance de la productivité aurait permis aux conseillers de prendre conscience du « bruit » du chronométrage de leur activité qui interfère avec leur prestation de service. Le fait d'expliciter les règles du contrôle de la fraude sociale ainsi que ses différentes conséquences aurait entraîné une prise de conscience des stratégies de contournement du contrôle qui contribuent à ne pas pérenniser l'activité commerciale de l'entreprise. Négocier les règles d'évaluation de la participation dans la formation aurait mis en perspective le caractère néfaste du « bachotage », c'est-à-dire l'effort d'apprentissage rapide d'une quantité importante de connaissances sur une période proche de l'examen, sur l'acquisition de compétences d'apprentissage durable tout au long de la vie.

La reconfiguration des dispositifs de vérification de l'information est donc passée par un élargissement des domaines d'opportunité des agents afin de réduire les inefficacités

organisationnelles en partie causées par l'existence de ces dispositifs. Le choix a ainsi été fait de donner davantage de pouvoir de décision à l'agent et non d'offrir cette opportunité au principal.

Les différents tests d'une potentielle amélioration de l'efficacité organisationnelle à la suite des modifications expérimentales du dispositif de vérification de l'information ont entraîné des résultats modestes, voire contre-intuitifs : la qualité de vie au travail des conseillers du centre d'appels a été détériorée, le consentement à la contribution sociale s'est infléchi vers moins de recrutement en CDD et la réussite des étudiants n'a pas été unidirectionnelle.

De manière générale, ces changements nous ont conduits à diagnostiquer l'existence de représentations individuelles du dispositif de vérification de l'information, c'est-à-dire des manières de percevoir son fonctionnement et son utilisation qui diffèrent de part et d'autre de la relation d'agence. Ces représentations ont été révélées par les expérimentations. Par conséquent, modifier le dispositif de vérification de l'information sans modifier ces représentations ne semble pas suffisant pour pouvoir améliorer l'efficacité organisationnelle.

Ce phénomène permet d'avancer l'idée selon laquelle le dispositif de vérification n'est pas neutre, c'est-à-dire qu'il n'est pas un outil mécanique dépourvu d'interprétations et de subjectivité et que son paramétrage dépend de la position des contractants dans la relation d'agence, de leurs interactions réciproques et de leurs attentes vis-à-vis du dispositif. Le dispositif de vérification est encastré dans un jeu de rôles et de représentations qui le dépassent et qu'il est difficile de modifier. Par conséquent, la réduction des inefficacités organisationnelles doit impliquer une remise en question de l'architecture globale de la relation d'agence et des dispositifs censés guider la coopération des individus. Le dispositif de vérification de l'information est un rouage important, mais il demeure à l'intersection d'autres considérations. Parmi celles-ci, on peut citer la confiance qui s'exprime entre les contractants, mais également à l'égard du dispositif de vérification (Ferrin et al., 2007).

Bien que la confiance intraorganisationnelle soit un concept ancien dans les sciences de gestion et en psychologie, en économie il ne semble pas être considéré avec autant d'importance, malgré des tentatives d'incorporation dans le champ de la théorie des contrats (Frey, 1993 ; Beccera et Gupta, 1999). Nous avons pu aborder l'idée que la confiance et la vérification de l'information étaient deux façons de régler le problème de la coopération. Dans ce cadre, il est possible d'envisager une complémentarité plutôt qu'une opposition entre ces deux processus. Par exemple, la possibilité offerte par le principal de donner une plus grande autonomie dans l'utilisation du dispositif de vérification de l'information à l'agent permettrait de donner une responsabilité plus importante dans la conduite de la tâche qui lui a été assignée. Cette confiance accordée dans l'utilisation du dispositif permettrait à l'agent de détecter lui-même ses erreurs et de les corriger. Néanmoins, la confiance expose les individus à un risque lorsque la vigilance sur la

rente informationnelle est baissée. L'un des contractants peut tirer parti du sentiment de sécurité par opportunisme. Dans ce cas et dans un cadre contractuel, il est donc également nécessaire de partager la prise en charge des risques liés à la confiance de la même façon que s'effectue le partage de l'assurance contre le risque engendré par un état du monde incertain.

D'autres considérations caractérisent le dispositif de surveillance, comme la réciprocité, la reconnaissance, ou encore l'équité de traitement qui sont autant de dispositions différentes de la récompense du travail rendu et qui environnent le champ d'action du principal et de l'agent. Ces concepts correspondent à la notion de contrat psychologique qui enveloppe les relations d'emploi (Conway et Briner, 2009). Il est sans doute possible d'adapter ces considérations à d'autres relations d'agence, mais ces réflexions vont bien au-delà de la problématique de la thèse. Il semble néanmoins nécessaire de les incorporer dans un champ d'études élargi afin de ne pas réduire la relation d'agence à seulement quatre dimensions : le principal, l'agent, le système de récompenses et le dispositif de vérification de l'information. Ces dimensions supplémentaires seront indispensables afin d'approfondir l'investigation dans ce domaine dans de futures recherches.

La principale recommandation que nous soumettons à l'issue de ces trois expérimentations sur l'économie des dispositifs de vérification de l'information est que leur paramétrage ainsi que leur motif d'utilisation doivent être explicites dès la mise en œuvre du contrat d'agence, au même titre que les règles de récompenses et de sanctions. L'agent doit être informé des motifs de l'utilisation du dispositif de vérification afin d'en apprécier la légitimité. *In fine*, incorporer le dispositif de vérification dans les termes du contrat permet de favoriser l'approbation de son utilisation. Le principal est ainsi contraint dans son pouvoir de vérification qui dépend alors de l'accord de l'agent.

III. Les principales limites de cette recherche

La principale limite de cette thèse de doctorat porte sur l'absence d'un support théorique général sur les dispositifs de vérification de l'information. En effet, chacun de nos chapitres utilisait un modèle spécifique à chaque relation étudiée. Pourtant, nous n'avons pas formulé un cadre théorique commun qui aurait permis d'interpréter nos résultats à l'aune de ces critères globaux. En l'absence de la formulation d'un modèle unifié, il n'est pas possible de confronter tous les résultats obtenus les uns vis-à-vis des autres.

Ce défaut de support théorique se mesure à l'aune des choix qui ont été entrepris pour tester les effets des variations du dispositif afin de résoudre des problèmes organisationnels. Dans ces conditions, nous n'avons pas testé de modèle structurel des dispositifs de vérification de l'information. Cette limite nous empêche directement de formuler des recommandations

supplémentaires de politique organisationnelle à portée générale, c'est-à-dire qu'à l'issue de nos travaux il est difficile de préconiser une configuration optimale du dispositif de vérification de l'information. Celle-ci dépend de paramètres que nous n'avons pas été en mesure de contrôler ou d'identifier avant la mise en place des expérimentations. En revanche, bien que ce travail de recommandation soit difficile, ces travaux ont pu mettre en lumière des éléments qui doivent être considérés dans la théorie.

La deuxième limite que nous souhaitons évoquer en conclusion tient à la méthode de recherche choisie pour étudier la résolution des problèmes organisationnels. Bien qu'elle ait permis de mesurer avec précision les effets de modifications d'environnement, la portée de l'expérimentation contrôlée par assignation aléatoire demeure restreinte par plusieurs facteurs. Les résultats sont locaux, ponctuels et dépendants des caractéristiques de l'échantillon qui est sélectionné. De ce fait, les interprétations que nous avons réalisées peuvent difficilement être généralisables à d'autres contextes : c'est le problème de la validité externe.

Ces lacunes peuvent être dépassées de plusieurs manières au cours de futurs travaux. Les protocoles expérimentaux peuvent être amendés sur ces trois terrains afin d'étendre leur validité au niveau spatial, temporel et individuel. Les problèmes de réception du traitement ou de contournement de la condition expérimentale peuvent être limités par une meilleure connaissance du terrain, par des relations plus étroites avec les parties prenantes et par l'instauration de traitements alternatifs.

Il serait également envisageable d'approfondir la mobilisation des protocoles mixtes de recherche qualitatif-quantitatif dans le cadre de prochaines expérimentations. Cet approfondissement passe par un développement des techniques d'enquête de terrain plus immersives, par exemple des monographies, ainsi que par un investissement dans des techniques de mesure quantitative, par exemple, dans des modules de questionnaire enrichis, qu'ils soient rétrospectifs ou longitudinaux. Associées, ces deux approches permettraient de donner à voir davantage de dimensions sous-jacentes aux modifications contrôlées.

Enfin, il serait possible de combiner les conditions d'expérimentation en laboratoire à celles du terrain de manière à promouvoir davantage de contrôle sur les tests effectués. Concilier les deux approches permettrait d'adopter des procédures de variations plus riches et d'atténuer les biais inhérents aux deux méthodes. Par exemple, le passage par le laboratoire permettrait d'étudier des représentations sur le dispositif de vérification de l'information dont le chercheur pourrait tenir compte une fois qu'il initialise une expérimentation de terrain.

IV. Les pistes ouvertes par cette recherche

L'approfondissement de cette thèse de doctorat passe d'abord par une formulation théorique plus précise des dispositifs de vérification de l'information qui incorporerait les pistes de modélisation que nous avons formulées et en utilisant les concepts que nous avons mobilisés dans des relations d'agence spécifiques : l'ambivalence des effets de la vérification ; la temporalité de l'exécution du contrat de vérification ; la crédibilité de l'explicitation des modalités de vérification ; la négociation des règles de la vérification ; l'intensité de la vérification.
La perspective d'un modèle plus unifié permet d'envisager d'autres paramètres de la vérification de l'information auxquels nous n'avons pas eu recours : le critère d'informativité du dispositif de vérification ; la position du dispositif, contingente à l'activité ou contingente à des indicateurs qui dépendent du comportement de l'agent ; la place des représentations ; les considérations de confiance et de réciprocité ; etc.

L'économie des dispositifs de vérification de l'information ne se restreint pas aux trois champs d'investigation dans lesquels nous nous sommes engagés. En effet, à partir du moment où se formalise une relation d'agence dans une organisation productive, le dispositif de vérification est présent, parce qu'il est à la base du système de récompenses / sanctions. Par exemple, le contrôle de la recherche d'emploi qui s'est répandu en Europe à la fin des années 2000 est un avatar de cette conceptualisation (Parent, 2014 ; Le Gall, 2015) : le demandeur d'emploi doit fournir des actes répétés de recherche qui sont contrôlés par les agents du service public pour l'emploi.
Dans le même ordre d'idée, et plus largement, le développement des démarches d'activation des politiques sociales qui prennent appui sur l'utilisation de la contractualisation entre les travailleurs sociaux et les bénéficiaires de prestations sociales afin d'associer un ensemble de devoirs sociaux à un ensemble de droits sociaux semble être un prolongement naturel de l'étude des règles de vérification de l'information. Dans ce contexte, il s'agit de contrôler le respect des engagements des allocataires des prestations sociales (Duvoux, 2006).

Une autre manière d'aborder l'économie des dispositifs de vérification de l'information serait de transposer ce cadre d'analyse à l'étude des nouvelles formes de création monétaire, les cryptomonnaies, et du « blockchain », qui sont apparus à la suite de la crise financière de 2007-2008. Dans ce cas particulier, la répartition des ressources, le financement de l'activité et les échanges via Internet ne font pas l'objet d'une vérification par une tierce partie, mais sont vérifiés par une liste « infalsifiable » de « preuves du travail » ou de « preuves de participation » fournies par des « mineurs ». Cette configuration suppose l'absence formelle du dispositif de vérification de l'information (Janin, 2018). Cette innovation technologique pose un certain nombre de

questions et fournit une piste supplémentaire pour intégrer de nouvelles dimensions à l'étude de la vérification de l'information, notamment celle de savoir dans quelle mesure la dilution du dispositif de vérification à l'ensemble de la communauté impose de nouvelles contraintes d'organisation de l'activité. Si tous les individus sont à la fois acteurs et vérificateurs, cela pose la question de la spécialisation et de la répartition des tâches.

Pour conclure, l'économie des dispositifs de vérification de l'information n'a pas encore montré toutes ses perspectives et leur importance devrait considérablement s'accroître au fur et à mesure de la maturation d'une organisation sociale décentralisée fonctionnant en réseau. La question de la coopération et de sa vérification dans les organisations n'est ainsi pas épuisée et demeure un champ d'investigations fertile.

Bibliographie

Abdelmalak, M. M. M. (2016). Faculty-Student Partnerships in Assessment. *International Journal of Teaching and Learning in Higher Education*, *28*(2), 193-203.

Adangnikou, N. (2008). Peut-on parler de recherche en pédagogie universitaire, aujourd'hui, en France ? *Revue des sciences de l'éducation*, *34*(3), 601-621. https://doi.org/10.7202/029510ar

Agence centrale des organismes de sécurité sociale. (2016). *Évaluation de l'évasion sociale*. Direction de la statistique, des études et de la prévision.

Aiello, J. R. (1993). Computer-Based Work Monitoring: Electronic Surveillance and Its Effects. *Journal of Applied Social Psychology*, *23*(7), 499-507. https://doi.org/10.1111/j.1559-1816.1993.tb01100.x

Aiello, J. R., & Shao, Y. (1993). Electronic performance monitoring and stress: The role of feedback and goal setting. In M. J. Smith & G. Salvendy (Éd.), *Human-Computer Interaction: Application and Case Studies* (p. 1011-1016). Amsterdam.

Aiello, J. R., & Svec, C. M. (1993). Computer Monitoring of Work Performance: Extending the Social Facilitation Framework to Electronic Presence. *Journal of Applied Social Psychology*, *23*(7), 537-548. https://doi.org/10.1111/j.1559-1816.1993.tb01102.x

Aina, C., Baici, E., Casalone, G., & Pastore, F. (2018). *The Economics of University Dropouts and Delayed Graduation: A Survey* (Discussion Paper No. 11421). IZA Institute of Labor Economics.

Ajzen, I., & Fishbein, M. (1977). Attitude-behavior relations: A theoretical analysis and review of empirical research. *Psychological Bulletin*, *84*(5), 888-918. https://doi.org/10.1037/0033-2909.84.5.888

Akerlof, G. A. (1970). The Market for « Lemons »: Quality Uncertainty and the Market Mechanism. *The Quarterly Journal of Economics*, *84*(3), 488-500. https://doi.org/10.2307/1879431

Akerlof, G. A., & Yellen, J. L. (1990). The Fair Wage-Effort Hypothesis and Unemployment. *The Quarterly Journal of Economics*, *105*(2), 255-283. https://doi.org/10.2307/2937787

Akşin, O. Z., de Véricourt, F., & Karaesmen, F. (2008). Call Center Outsourcing Contract Analysis and Choice. *Management Science*, *54*(2), 354-368. https://doi.org/10.1287/mnsc.1070.0823

Alchian, A. A., & Demsetz, H. (1972). Production, Information Costs, and Economic Organization. *The American Economic Review*, *62*(5), 777–795.

Alder, G. S., & Ambrose, M. L. (2005). An examination of the effect of computerized performance monitoring feedback on monitoring fairness, performance, and satisfaction. *Organizational Behavior and Human Decision Processes*, *97*(2), 161-177. https://doi.org/10.1016/j.obhdp.2005.03.003

Algava, E., & Amira, S. (2011). *Sous-traitance : des relations de travail plus difficile chez les preneurs d'ordre* (DARES Analyse). Direction de l'animation de la recherche, des études et des statistiques.

Alge, B. J., Ballinger, G. A., & Green, S. G. (2004). Remote control: predictors of electronic monitoring intensity and secrecy. *Personnel Psychology*, *57*(2), 377-410. https://doi.org/10.1111/j.1744-6570.2004.tb02495.x

Alge, B. J., & Hansen, S. (2014). Workplace monitoring and surveillance research since 1984: A review and agenda. In M. D. Coovert & L. F. Thompson (Éd.), *The psychology of workplace technology* (p. 209-237). New York London: Routledge, Taylor & Francis Group.

Allingham, M. G., & Sandmo, A. (1972). Income tax evasion: a theoretical analysis. *Journal of Public Economics*, *1*(3-4), 323-338. https://doi.org/10.1016/0047-2727(72)90010-2

Alm, J. (2012). Measuring, explaining, and controlling tax evasion: lessons from theory, experiments, and field studies. *International Tax and Public Finance*, *19*(1), 54-77. https://doi.org/10.1007/s10797-011-9171-2

Alm, J. (2018). What motivates tax compliance? *Journal of Economic Surveys*, (forthcoming), 1-36. https://doi.org/10.1111/joes.12272

Alm, J., & McClellan, C. (2012). Tax Morale and Tax Compliance from the Firm's Perspective. *Kyklos*, *65*(1), 1-17. https://doi.org/10.1111/j.1467-6435.2011.00524.x

Andreoni, J., Erard, B., & Feinstein, J. (1998). Tax Compliance. *Journal of Economic Literature*, *36*(2), 818-860.

Angrist, J. D., & Pischke, J.-S. (2010). The Credibility Revolution in Empirical Economics: How Better Research Design is Taking the Con out of Econometrics. *Journal of Economic Perspectives*, *24*(2), 3-30. https://doi.org/10.1257/jep.24.2.3

Arnaud, S., & Chandon, J.-L. (2013). Will monitoring systems kill intrinsic motivation? An empirical study. *Revue de Gestion Des Ressources Humaines*, *90*(4), 35-53. https://doi.org/10.3917/grhu.090.0035

Arnold, I. J. M., & Straten, J. T. (2012). Motivation and Math Skills as Determinants of First-Year Performance in Economics. *The Journal of Economic Education*, *43*(1), 33-47. https://doi.org/10.1080/00220485.2012.636709

Arrow, K. J. (1962). The Economic Implications of Learning by Doing. *The Review of Economic Studies*, *29*(3), 155-173. https://doi.org/10.2307/2295952

Arrow, K. J. (1970). The organization of economic activity: issues pertinent to the choice of market versus nonmarket allocation. In R. H. Haveman & J. Margolis (Éd.), *Public expenditures and policy analysis* (p. 59-73). Chicago: Markham Pub. Co.

Ashenfelter, O. (1987). The case for evaluating training programs with randomized trials. *Economics of Education Review*, *6*(4), 333-338. https://doi.org/10.1016/0272-7757(87)90016-1

Auerbach, A. J., & Kotlikoff, L. J. (1987). Evaluating Fiscal Policy with a Dynamic Simulation Model. *The American Economic Review*, *77*(2), 49-55.

Azariadis, C. (1975). Implicit Contracts and Underemployment Equilibria. *Journal of Political Economy*, *83*(6), 1183-1202. https://doi.org/10.1086/260388

Bailey, C., & Turner, J. (2001). Strategies to Reduce Contribution Evasion in Social Security Financing. *World Development*, *29*(2), 385-393. https://doi.org/10.1016/S0305-750X(00)00100-5

Baiman, S., & Demski, J. S. (1980). Economically Optimal Performance Evaluation and Control Systems. *Journal of Accounting Research*, *18*, 184-220. https://doi.org/10.2307/2490338

Bain, P., & Taylor, P. (2000). Entrapped by the 'electronic panopticon'? Worker resistance in the call centre. *New Technology, Work and Employment*, *15*(1), 2-18. https://doi.org/10.1111/1468-005X.00061

Bain, P., Watson, A., Mulvey, G., Taylor, P., & Gall, G. (2002). Taylorism, targets and the pursuit of quantity and quality by call centre management. *New Technology, Work and Employment*, *17*(3), 170-185. https://doi.org/10.1111/1468-005X.00103

Baker, G. P., Jensen, M. C., & Murphy, K. J. (1988). Compensation and Incentives: Practice vs. Theory. *The Journal of Finance*, *43*(3), 593-616. https://doi.org/10.1111/j.1540-6261.1988.tb04593.x

Baktavatsalou, R., Maury-Duprey, D., Rathle, J.-P., & Tagnani, S. (2015). *Les entreprises en France - Édition 2015* (Insee Références). Paris: Institut national de la statistique et des études économiques. Consulté à l'adresse https://www.insee.fr/fr/statistiques/1906539

Balkenborg, D. (2001). How Liable Should a Lender Be? The Case of Judgment-Proof Firms and Environmental Risk: Comment. *American Economic Review*, *91*(3), 731-738. https://doi.org/10.1257/aer.91.3.731

Ball, K. S. (2010). Workplace surveillance: an overview. *Labor History*, *51*(1), 87-106. https://doi.org/10.1080/00236561003654776

Ball, K. S., & Margulis, S. T. (2011). Electronic monitoring and surveillance in call centres: a framework for investigation. *New Technology, Work and Employment*, *26*(2), 113-126. https://doi.org/10.1111/j.1468-005X.2011.00263.x

Bandiera, O., Barankay, I., & Rasul, I. (2011). Field Experiments with Firms. *Journal of Economic Perspectives*, *25*(3), 63-82. https://doi.org/10.1257/jep.25.3.63

Banerjee, A., Duflo, E., Glennerster, R., & Kinnan, C. (2015). The Miracle of Microfinance? Evidence from a Randomized Evaluation. *American Economic Journal: Applied Economics*, *7*(1), 22-53. https://doi.org/10.1257/app.20130533

Barak, A. (2012). Deterrence and moral persuasion effects on corporate tax compliance: findings from a randomized controlled trial. *Criminology*, *50*(1), 27-69. https://doi.org/10.1111/j.1745-9125.2011.00256.x

Barkema, H. G. (1995). Do Top Managers Work Harder When They Are Monitored? *Kyklos*, *48*(1), 19-42. https://doi.org/10.1111/j.1467-6435.1995.tb02313.x

Barnard, C. I., & Andrews, K. R. (1938). *The functions of the executive*. Cambridge, MA: Harvard Univ. Press.

Bartels, L. K., & Nordstrom, C. R. (2012). Examining big brother's purpose for using electronic performance monitoring. *Performance Improvement Quarterly*, *25*(2), 65-77. https://doi.org/10.1002/piq.20140

Bates, R. A., & Holton, E. F. (1995). Computerized performance monitoring: A review of human resource issues. *Human Resource Management Review*, *5*(4), 267-288. https://doi.org/10.1016/1053-4822(95)90010-1

Batt, R. L., Doellgast, V., & Kwon, H. (2005). Service Management and Employment Systems in U.S. and Indian Call Centers. *Brookings Trade Forum*, *2005*(1), 335-360. https://doi.org/10.1353/btf.2006.0004

Baumann, F., Friehe, T., & Jansen, M. (2009). On the Economics of Contribution Evasion. *FinanzArchiv: Public Finance Analysis*, *65*(2), 162-177. https://doi.org/10.1628/001522109X466509

Bayer, R., & Cowell, F. (2009). Tax compliance and firms' strategic interdependence. *Journal of Public Economics*, *93*(11-12), 1131-1143. https://doi.org/10.1016/j.jpubeco.2009.07.007

Beccera, M., & Gupta, A. K. (1999). Trust within the organization: Integrating the trust literature with agency theory and transaction costs economics. *Public Administration Quarterly*, *23*(2), 177-203.

Becker, G. S. (1962). Investment in Human Capital: A Theoretical Analysis. *Journal of Political Economy*, *70*(5, Part 2), 9-49. https://doi.org/10.1086/258724

Becker, G. S. (1968). Crime and Punishment: An Economic Approach. *Journal of Political Economy*, *76*(2), 169-217. https://doi.org/10.1086/259394

Becker, W. E., & Rosen, S. (1992). The learning effect of assessment and evaluation in high school. *Economics of Education Review*, *11*(2), 107-118. https://doi.org/10.1016/0272-7757(92)90002-K

Beffy, M., Fougère, D., & Maurel, A. (2009). L'impact du travail salarié des étudiants sur la réussite et la poursuite des études universitaires. *Économie et Statistique*, *422*(1), 31-50. https://doi.org/10.3406/estat.2009.8017

Bellity, E., Gilles, F., & L'Horty, Y. (2017). *Does practicing literacy skills improve academic performance in first-year university students? Results from a randomized experiment* (Working Paper No. 2017-02). Travail, emploi et politiques publiques.

Belot, M., & Schröder, M. (2015). The Spillover Effects of Monitoring: A Field Experiment. *Management Science*, *62*(1), 37-45. https://doi.org/10.1287/mnsc.2014.2089

Ben Fekih Aissi, L., & Naudin, M. (2015). Les stratégies d'ajustement de téléopérateurs de centres d'appels face au stress généré par le monitoring électronique des performances. *La Revue des Sciences de Gestion*, *275-276*(5), 85-97. https://doi.org/10.3917/rsg.275.0085

Bénabou, R., & Tirole, J. (2003). Intrinsic and Extrinsic Motivation. *Review of Economic Studies*, *70*(3), 489-520. https://doi.org/10.1111/1467-937X.00253

Bénabou, R., & Tirole, J. (2006). Incentives and Prosocial Behavior. *American Economic Review*, *96*(5), 1652-1678. https://doi.org/10.1257/aer.96.5.1652

Bental, B., Deffains, B., & Demougin, D. (2012). Credibility and Monitoring: Outsourcing as a Commitment Device. *Journal of Economics & Management Strategy*, *21*(1), 31-52. https://doi.org/10.1111/j.1530-9134.2011.00327.x

Bertrand, C. (2014). *Soutenir la transformation pédagogique dans l'enseignement supérieur*. Ministère de l'Enseignement Supérieur et de la Recherche.

Bhave, D. P. (2013). The Invisible Eye? Electronic Performance Monitoring and Employee Job Performance. *Personnel Psychology*, *67*(3), 605-635. https://doi.org/10.1111/peps.12046

Birenbaum, M. (1996). Assessment 2000: Towards a Pluralistic Approach to Assessment. In M. Birenbaum & F. J. R. C. Dochy (Éd.), *Alternatives in Assessment of Achievements, Learning Processes and Prior Knowledge* (p. 3-29). Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-011-0657-3_1

Bishop, J. (2006). Chapter 15: Drinking from the Fountain of Knowledge: Student Incentive to Study and Learn – Externalities, Information Problems and Peer Pressure. In *Handbook of the Economics of Education* (Vol. 2, p. 909-944). Elsevier. https://doi.org/10.1016/S1574-0692(06)02015-0

Black, P., & Wiliam, D. (1998). Assessment and Classroom Learning. *Assessment in Education: Principles, Policy & Practice*, *5*(1), 7-74. https://doi.org/10.1080/0969595980050102

Blackwell, D. (1953). Equivalent Comparisons of Experiments. *The Annals of Mathematical Statistics*, *24*(2), 265-272.

Bloom, B. S., Hastings, T. J., & Madaus, G. F. (1971). *Handbook on formative and summative evaluation of student learning*. New York: McGraw-Hill.

Bloom, H. S. (2008). Chapter 9: The Core Analytics of Randomized Experiments for Social Research. In P. Alasuutari, L. Bickman, & J. Brannen (Éd.), *The SAGE Handbook of social research methods* (p. 115-133). Los Angeles: Sage.

Bloxham, S., den-Outer, B., Hudson, J., & Price, M. (2016). Let's stop the pretence of consistent marking: exploring the multiple limitations of assessment criteria. *Assessment & Evaluation in Higher Education*, *41*(3), 466-481. https://doi.org/10.1080/02602938.2015.1024607

Bolton, P., Dewatripont, M., & Campbell, A. (2005). *Contract theory*. Cambridge, Mass: MIT Press.

Boly, A. (2011). On the incentive effects of monitoring: evidence from the lab and the field. *Experimental Economics*, *14*(2), 241-253. https://doi.org/10.1007/s10683-010-9265-1

Bonesrønning, H. (2004). Do the teachers' grading practices affect student achievement? *Education Economics*, *12*(2), 151-167. https://doi.org/10.1080/0964529042000239168

Bonesrønning, H., & Opstad, L. (2015). Can student effort be manipulated? Does it matter? *Applied Economics*, *47*(15), 1511-1524. https://doi.org/10.1080/00036846.2014.997923

Boud, D., & Falchikov, N. (Éd.). (2007). *Rethinking assessment in higher education: learning for the longer term*. Routledge.

Bourgeois, L. (2008). *Solidarité : l'idée de solidarité et ses conséquences sociales*. Le Bord de l'eau.

Bouvier, O.-L. (2007). Le travail dissimulé. *Informations Sociales*, *142*, 120-131.

Bovill, C., Cook- Sather, A., & Felten, P. (2011). Students as co- creators of teaching approaches, course design, and curricula: implications for academic developers. *International Journal for Academic Development*, *16*(2), 133-145. https://doi.org/10.1080/1360144X.2011.568690

Brehm, J. W. (1966). *A theory of psychological reactance*. Oxford, England: Academic Press.

Brousseau, E., & Glachant, J.-M. (2000). Introduction : Économie des contrats et renouvellements de l'analyse économique. *Revue d'économie industrielle*, *92*(1), 23-50. https://doi.org/10.3406/rei.2000.1033

Brown, B. W., & Saks, D. H. (1981). The Microeconomics of Schooling. *Review of Research in Education*, *9*(1), 217-254. https://doi.org/10.3102/0091732X009001217

Brown, B. W., & Saks, D. H. (1987). The microeconomics of the allocation of teachers' time and student learning. *Economics of Education Review*, *6*(4), 319-332. https://doi.org/10.1016/0272-7757(87)90015-X

Brown, S., & Knight, P. (1994). *Assessing learners in higher education*. London ; Philadelphia: Kogan Page.

Burtless, G. (1995). The Case for Randomized Field Trials in Economic and Policy Research. *Journal of Economic Perspectives*, *9*(2), 63-84. https://doi.org/10.1257/jep.9.2.63

Buscatto, M. (2002). Les centres d'appels, usines modernes ? Les rationalisations paradoxales de la relation téléphonique. *Sociologie du Travail*, *44*(1), 99-117. https://doi.org/10.1016/S0038-0296(01)01202-X

Callaghan, G., & Thompson, P. (2001). Edwards Revisited: Technical Control and Call Centres. *Economic and Industrial Democracy*, *22*(1), 13-37. https://doi.org/10.1177/0143831X01221002

Calvet Christian, R., & Alm, J. (2014). Empathy, sympathy, and tax compliance. *Journal of Economic Psychology*, *40*, 62-82. https://doi.org/10.1016/j.joep.2012.10.001

Calvo, G. A., & Wellisz, S. (1978). Supervision, Loss of Control, and the Optimum Size of the Firm. *Journal of Political Economy*, *86*(5), 943-952.

Camerer, C., Loewenstein, G., & Rabin, M. (Éd.). (2004). *Advances in behavioral economics*. New York: Princeton, N.J: Russell Sage Foundation ; Princeton University Press.

Carayon, P. (1993). Effect of Electronic Performance Monitoring on Job Design and Worker Stress: Review of the Literature and Conceptual Model. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *35*(3), 385-395. https://doi.org/10.1177/001872089303500301

Card, D. (1999). Chapter 30: The Causal Effect of Education on Earnings. In *Handbook of Labor Economics* (Vol. 3, p. 1801-1863). Elsevier. https://doi.org/10.1016/S1573-4463(99)03011-4

Chalykoff, J., & Kochan, T. A. (1989). Computer-aided monitoring: its influence on employee job satisfaction and turnover. *Personnel Psychology*, *42*(4), 807-834. https://doi.org/10.1111/j.1744-6570.1989.tb00676.x

Chanson, G. (2003). Externalisation et performance dans la relation client-prestataire. *Revue française de gestion*, *29*(143), 43-54. https://doi.org/10.3166/rfg.143.43-54

Chaulet, J. (2006). Parler et faire parler : travail en centre d'appels et dispositif de surveillance équipée. *Terrains & travaux*, *11*(2), 36-60.

Chen, K.-P., & Chu, C. Y. C. (2005). Internal Control versus External Manipulation: A Model of Corporate Income Tax Evasion. *The RAND Journal of Economics*, *36*(1), 151-164.

Chen, R., & Sanders, G. L. (2007). Electronic Monitoring in Workplace: Synthesis and Theorizing. In *AMCIS 2007 Proceedings* (p. 1-9).

Chouanière, D., Boini, S., & Colin, R. (2011). Conditions de travail et santé dans les centres d'appels téléphoniques. *Documents pour le médecin du travail*, (126), 241-259.

Clotfelter, C. T. (1983). Tax Evasion and Tax Rates: An Analysis of Individual Returns. *The Review of Economics and Statistics*, *65*(3), 363-373. https://doi.org/10.2307/1924181

Coase, R. H. (1937). The Nature of the Firm. *Economica*, *4*(16), 386-405. https://doi.org/10.1111/j.1468-0335.1937.tb00002.x

Congdon, W. J., Kling, J. R., & Mullainathan, S. (2011). *Policy and choice: public finance through the lens of behavioral economics*. Washington, D.C: Brookings Institution Press.

Connell, J., & Hannif, Z. (2009). Call centres, quality of work life and HRM practices: An in-house/outsourced comparison. *Employee Relations*, *31*(4), 363-381. https://doi.org/10.1108/01425450910965423

Conseil des Prélèvements Obligatoires. (2007). *La fraude aux prélèvements obligatoires et son contrôle*.

Conway, N., & Briner, R. B. (2009). Fifty years of psychological contract research: What do we know and what are the main challenges. *International review of industrial and organizational psychology*, *24*(71), 71-131.

Correa, H., & Gruver, G. W. (1987). Teacher-student interaction: A game theoretic extension of the economic theory of education. *Mathematical Social Sciences*, *13*(1), 19-47. https://doi.org/10.1016/0165-4896(87)90057-6

Costrell, R. M. (1994). A Simple Model of Educational Standards. *American Economic Review*, *84*(4), 956-971.

Cousin, O. (2002). Les ambivalences du travail Les salariés peu qualifiés dans les centres d'appels. *Sociologie du Travail*, *44*(4), 499-520. https://doi.org/10.1016/S0038-0296(02)01259-1

Cowen, T., & Glazer, A. (1996). More monitoring can induce less effort. *Journal of Economic Behavior & Organization*, *30*(1), 113-123. https://doi.org/10.1016/S0167-2681(96)00845-1

Crocker, K. J., & Slemrod, J. (2005). Corporate tax evasion with agency costs. *Journal of Public Economics*, *89*(9-10), 1593-1610. https://doi.org/10.1016/j.jpubeco.2004.08.003

Cummings, R. G., Martinez-Vazquez, J., McKee, M., & Torgler, B. (2009). Tax morale affects tax compliance: Evidence from surveys and an artefactual field experiment. *Journal of Economic Behavior & Organization*, *70*(3), 447-457. https://doi.org/10.1016/j.jebo.2008.02.010

Davidson, R., & Henderson, R. (2000). Electronic Performance Monitoring: A Laboratory Investigation of the Influence of Monitoring and Difficulty on Task Performance, Mood State, and Self-Reported Stress Levels. *Journal of Applied Social Psychology*, *30*(5), 906-920. https://doi.org/10.1111/j.1559-1816.2000.tb02502.x

Davis, L. E., North, D. C., & Smorodin, C. (1971). *Institutional change and American economic growth*. Cambridge: University Press.

Davis, S. J., & Haltiwanger, J. (1999). Chapter 41: Gross job flows. In *Handbook of Labor Economics* (Vol. 3, p. 2711-2805). Elsevier. https://doi.org/10.1016/S1573-4463(99)30027-4

De Bandt, J. (1988). Le débat sur la productivité dans les services : des problèmes mal posés. *Revue d'économie industrielle*, *43*(1), 179-195. https://doi.org/10.3406/rei.1988.1017

Deaton, A. (2010). Instruments, Randomization, and Learning about Development. *Journal of Economic Literature*, *48*(2), 424-455. https://doi.org/10.1257/jel.48.2.424

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York: Plenum.

Deci, E. L., Spiegel, N. H., Koestner, R., & Kauffman, M. (1982). Effects of performance standards on teaching styles: Behavior of controlling teachers. *Journal of Educational Psychology*, *74*(6), 852-859. https://doi.org/10.1037/0022-0663.74.6.852

Deeley, S. J. (2014). Summative co-assessment: A deep learning approach to enhancing employability skills and attributes. *Active Learning in Higher Education*, *15*(1), 39-51. https://doi.org/10.1177/1469787413514649

Deeley, S. J., & Bovill, C. (2017). Staff student partnership in assessment: enhancing assessment literacy through democratic practices. *Assessment & Evaluation in Higher Education*, *42*(3), 463-477. https://doi.org/10.1080/02602938.2015.1126551

Deeley, S. J., & Brown, R. A. (2014). *Learning Through Partnership in Assessment* (Teaching and learning together in higher education No. 13). Bryn Mawr College.

Demarco, G., & Rofman, R. (1999). *Collecting and Transferring Pension Contributions* (Working Paper No. 20122). Banque Mondiale.

Demougin, D., & Fluet, C. (2001). Monitoring versus incentives. *European Economic Review*, *45*(9), 1741-1764. https://doi.org/10.1016/S0014-2921(00)00051-9

Demougin, D., & Helm, C. (2006). Moral Hazard and Bargaining Power. *German Economic Review*, *7*(4), 463-470. https://doi.org/10.1111/j.1468-0475.2006.00130.x

Di Porto, E. (2011). Undeclared Work, Employer Tax Compliance, and Audits. *Public Finance Review*, *39*(1), 75-102. https://doi.org/10.1177/1091142110381641

Diaye, M.-A., & Koshevoy, G. (2006). *Black Market, Labor Demand and Tax Evasion*. Centre d'Études des Politiques Économiques (EPEE), Université d'Evry Val d'Essonne. Consulté à l'adresse https://ideas.repec.org/p/eve/wpaper/06-04.html

Dickinson, D., & Villeval, M.-C. (2008). Does monitoring decrease work effort? *Games and Economic Behavior*, *63*(1), 56-76. https://doi.org/10.1016/j.geb.2007.08.004

Dittrich, M., & Städter, S. (2015). Moral hazard and bargaining over incentive contracts. *Research in Economics*, *69*(1), 75-85. https://doi.org/10.1016/j.rie.2014.10.002

Djellal, F., & Gallouj, F. (2007). Les services publics à l'épreuve de la productivité et la productivité à l'épreuve des services publics. *Revue d'économie industrielle*, (119), 25-54. https://doi.org/10.4000/rei.1963

Dochy, F., Segers, M., & Sluijsmans, D. (1999). The use of self-, peer and co-assessment in higher education: A review. *Studies in Higher Education*, *24*(3), 331-350. https://doi.org/10.1080/03075079912331379935

Doerrenberg, P., & Schmitz, J. (2015). Tax Compliance and Information Provision A Field Experiment with Small Firms. https://doi.org/10.2139/ssrn.2608118

Douthitt, E. A., & Aiello, J. R. (2001). The role of participation and control in the effects of computer monitoring on fairness perceptions, task satisfaction, and performance. *Journal of Applied Psychology*, *86*(5), 867-874. https://doi.org/10.1037//0021-9010.86.5.867

Doyle, E., Buckley, P., & Whelan, J. (2018). Assessment co-creation: an exploratory analysis of opportunities and challenges based on student and instructor perspectives. *Teaching in Higher Education*, 1-16. https://doi.org/10.1080/13562517.2018.1498077

Duguet, A., Le Mener, M., & Morlaix, S. (2016). The Key Predictors of Success in University in France: What Are the Contributing Factors and Possible New Directions in Educational Research? *International Journal of Higher Education*, *5*(3), 222-235. https://doi.org/10.5430/ijhe.v5n3p222

Duguet, A., & Morlaix, S. (2012). Les pratiques pédagogiques des enseignants universitaires : Quelle variété pour quelle efficacité ? *Questions vives recherches en éducation*, *6*(18), 93-110. https://doi.org/10.4000/questionsvives.1178

Duvoux, N. (2007). 22. Le rmi et les dérives de la contractualisation. In S. Paugam, *Repenser la solidarité* (p. 451-472). Presses Universitaires de France. https://doi.org/10.3917/puf.pauga.2007.02.0451

Dye, R. A. (1986). Optimal Monitoring Policies in Agencies. *The RAND Journal of Economics*, *17*(3), 339-350. https://doi.org/10.2307/2555715

Earley, C. P. (1988). Computer-generated performance feedback in the magazine-subscription industry. *Organizational Behavior and Human Decision Processes*, *41*(1), 50-64. https://doi.org/10.1016/0749-5978(88)90046-5

Eeckhoudt, L., Gollier, C., & Schlesinger, H. (2005). *Economic and financial decisions under risk*. Princeton, N.J: Princeton University Press.

Eisenhardt, K. M. (1989). Agency Theory: An Assessment and Review. *Academy of Management Review*, *14*(1), 57-74. https://doi.org/10.5465/amr.1989.4279003

Ellingsen, T., & Johannesson, M. (2008). Pride and Prejudice: The Human Side of Incentive Theory. *American Economic Review*, *98*(3), 990-1008. https://doi.org/10.1257/aer.98.3.990

Enoff, L. D., & McKinnon, R. (2011). Recouvrement des cotisations et respect des obligations : améliorer la gouvernance pour élargir la couverture sociale. *Revue internationale de sécurité sociale*, *64*(4), 111-133. https://doi.org/10.1111/j.1752-1718.2011.01413.x

Fagart, M.-C., & Sinclair-Desgagné, B. (2007). Ranking Contingent Monitoring Systems. *Management Science*, *53*(9), 1501-1509. https://doi.org/10.1287/mnsc.1060.0693

Falchikov, N. (1986). Product comparisons and process benefits of collaborative peer group and self-assessments. *Assessment & Evaluation in Higher Education*, *11*(2), 146-166. https://doi.org/10.1080/0260293860110206

Falk, A., & Koslfeld, M. (2006). The Hidden Costs of Control. *The American Economic Review*, *96*(5), 1611-1630.

Fama, E. F., & Jensen, M. C. (1983). Separation of Ownership and Control. *The Journal of Law and Economics*, *26*(2), 301-325. https://doi.org/10.1086/467037

Fehr, E., & Falk, A. (2002). Psychological foundations of incentives. *European Economic Review*, *46*(4-5), 687-724. https://doi.org/10.1016/S0014-2921(01)00208-2

Feld, L. P., & Frey, B. S. (2002). Trust breeds trust: How taxpayers are treated. *Economics of Governance*, *3*(2), 87-99. https://doi.org/10.1007/s101010100032

Fellner, G., Sausgruber, R., & Traxler, C. (2013). Testing Enforcement Strategies in the Field. *Journal of the European Economic Association*, *11*(3), 634-660. https://doi.org/10.1111/jeea.12013

Ferber, R., & Hirsch, W. Z. (1982). *Social experimentation and economic policy*. Cambridge ; New York: Cambridge University Press.

Ferrin, D. L., Bligh, M. C., & Kohles, J. C. (2007). Can I Trust You to Trust Me? A Theory of Trust, Monitoring, and Cooperation in Interpersonal and Intergroup Relationships. *Group & Organization Management*, *32*(4), 465-499. https://doi.org/10.1177/1059601106293960

Festré, A., & Garrouste, P. (2015). Theory and evidence in psychology and economics about motivation crowding out: a possible convergence? *Journal of Economic Surveys*, *29*(2), 339-356. https://doi.org/10.1111/joes.12059

Fimbel, E. (2003). Nature enjeux et stratégiques de l'externalisation. *Revue française de gestion*, *29*(143), 27-42. https://doi.org/10.3166/rfg.143.27-42

Findlay, P., Kalleberg, A. L., & Warhurst, C. (2013). The challenge of job quality. *Human Relations*, *66*(4), 441-451. http://journals.sagepub.com/doi/abs/10.1177/0018726713481070

Fischer, R. (1935). *The design of experiments*. Edinburgh: Oliver and Boyd.

Foucault, M. (1975). *Surveiller et punir : naissance de la prison*. Paris: Gallimard.

Fougère, D. (2012). Les méthodes d'expérimentation en question. *Éducation et formations*, (81), 41-48.

Frey, B. S. (1993). Does monitoring increase work effort? The rivalry with trust and loyalty. *Economic Inquiry*, *31*(4), 663-670. https://doi.org/10.1111/j.1465-7295.1993.tb00897.x

Frey, B. S., & Jegen, R. (2001). Motivation Crowding Theory. *Journal of Economic Surveys*, *15*(5), 589-611. https://doi.org/10.1111/1467-6419.00150

Fugazza, M., & Jacques, J.-F. (2004). Labor market institutions, taxation and the underground economy. *Journal of Public Economics*, *88*(1-2), 395-418. https://doi.org/10.1016/S0047-2727(02)00079-8

Gagné, M., & Deci, E. L. (2005). Self-determination theory and work motivation. *Journal of Organizational Behavior*, *26*(4), 331-362. https://doi.org/10.1002/job.322

Gaignon, A., & Lieurade, C. (2018). *Le contrôle de la recherche d'emploi : l'impact sur le parcours des demandeurs d'emploi* (Éclairages et Synthèse No. 48). Pôle Emploi - Statistiques, études et évaluations.

Gangl, K., Torgler, B., Kirchler, E., & Hofmann, E. (2014). Effects of supervision on tax compliance: Evidence from a field experiment in Austria. *Economics Letters*, *123*(3), 378-382. https://doi.org/10.1016/j.econlet.2014.03.027

Gans, N., & Zhou, Y.-P. (2007). Call-Routing Schemes for Call-Center Outsourcing. *Manufacturing & Service Operations Management*, *9*(1), 33-50. https://doi.org/10.1287/msom.1060.0119

Garside, J., Nhemachena, J. Z. Z., Williams, J., & Topping, A. (2009). Repositioning assessment: Giving students the 'choice' of assessment methods. *Nurse Education in Practice*, *9*(2), 141-148. https://doi.org/10.1016/j.nepr.2008.09.003

Gausson, L. (2017). *Plus de la moitié des nouveaux bacheliers passent en deuxième année* (Ofipe résultats No. 159). Observatoire des formations, des Insertions Professionnelles, Évaluations.

Gausson, L. (2018). *Plus de 12 000 étudiants en 2017-2018* (Ofipe résultats No. 162). Observatoire des formations, des Insertions Professionnelles, Évaluations.

Gauthier, R.-F., Caffin-Ravier, M., Descamps, B., Mosnier, M., & Peretti, H. (2007). *L'évaluation des étudiants à l'Université : point aveugle ou point d'appui ?* (No. 2007-072). Inspection générale de l'administration de l'Éducation nationale et de la Recherche.

Gibbons, M. (Éd.). (1994). *The new production of knowledge: the dynamics of science and research in contemporary societies*. London: SAGE Publications.

Gibbons, R., & Roberts, J. (Éd.). (2013). *The handbook of organizational economics*. Princeton: Princeton University Press.

Gibbs, G., & Simpson, C. (2004). Conditions Under Which Assessment Supports Students' Learning. *Learning and Teaching in Higher Education*, *1*, 3-31.

Gillet, I., Greenan, N., & Le Gall, R. (2016). *Les effets de la surveillance électronique : une expérimentation dans un centre d'appels*. Paris: Lavoisier.

Gillion, C., Turner, J. A., Bailey, C., & Latulippe, D. (Éd.). (2000). *Social security pensions: development and reform*. Geneva: International Labour Office.

Gjesdal, F. (1982). Information and Incentives: The Agency Information Problem. *The Review of Economic Studies*, *49*(3), 373-390. https://doi.org/10.2307/2297362

Gneezy, U., Meier, S., & Rey-Biel, P. (2011). When and Why Incentives (Don't) Work to Modify Behavior. *Journal of Economic Perspectives*, *25*(4), 191-210. https://doi.org/10.1257/jep.25.4.191

Goerke, L. (2014). Tax Evasion by Firms. In J. Backhaus (Éd.), *Encyclopedia of Law and Economics* (p. 1-7). New York, NY: Springer New York. https://doi.org/10.1007/978-1-4614-7883-6_34-1

Goerke, L., & Runkel, M. (2006). Profit Tax Evasion under Oligopoly with Endogenous Market Structure. *National Tax Journal*, *59*(4), 851-857.

Gonzales, L., Roussel, R., Héam, J.-C., Mikou, M., & Ferretti, C. (2018). *La protection sociale en France et en Europe en 2016 - Résultats des comptes de la protection sociale*. Direction de la recherche, des études, de l'évaluation et des statistiques (DRESS). Consulté à l'adresse http://drees.solidarites-sante.gouv.fr/etudes-et-statistiques/publications/panoramas-de-la-drees/article/la-protection-sociale-en-france-et-en-europe-en-2016-resultats-des-comptes-de

Goveia, L., & Sosa, A. (2017). Developing a compliance-based approach to address error, evasion and fraud in social security systems: A compliance-based approach to EEF. *International Social Security Review*, *70*(2), 87-107. https://doi.org/10.1111/issr.12136

Gownder, J. P., Voce, C., Schadler, T., Ask, J. A., Thayer, F., & Williamson, K. (2014). *The Enterprise Wearables Journey: How Wearable Computing Will Improve Your Business*. Forrester Research.

Grant, R., & Higgins, C. (1989). Monitoring service workers via computer: The effect on employees, productivity, and service. *National Productivity Review*, *8*(2), 101-112. https://doi.org/10.1002/npr.4040080203

Grebner, S., Semmer, N., Faso, L. L., Gut, S., Kälin, W., & Elfering, A. (2003). Working conditions, well-being, and job-related attitudes among call centre agents. *European Journal of Work and Organizational Psychology*, *12*(4), 341-365. https://doi.org/10.1080/13594320344000192

Griffith, T. L. (1993). Monitoring and Performance: A Comparison of Computer and Supervisor Monitoring1. *Journal of Applied Social Psychology*, *23*(7), 549-572. https://doi.org/10.1111/j.1559-1816.1993.tb01103.x

Grosjean, V., & Ribert-Van De Weerdt, C. (2005). Vers une psychologie ergonomique du bien-être et des émotions : les effets du contrôle dans les centres d'appels. *Le travail humain*, *68*(4), 355-378. https://doi.org/10.3917/th.684.0355

Grossman, S. J., & Hart, O. D. (1983). An Analysis of the Principal-Agent Problem. *Econometrica*, *51*(1), 7-45. https://doi.org/10.2307/1912246

Handelsman, M. M., Briggs, W. L., Sullivan, N., & Towler, A. (2005). A Measure of College Student Course Engagement. *The Journal of Educational Research*, *98*(3), 184-192. https://doi.org/10.3200/JOER.98.3.184-192

Harrison, G. W. (2013). Field experiments and methodological intolerance. *Journal of Economic Methodology*, *20*(2), 103-117. https://doi.org/10.1080/1350178X.2013.804678

Harrison, G. W., & List, J. A. (2004). Field Experiments. *Journal of Economic Literature*, *42*(4), 1009-1055. https://doi.org/10.1257/0022051043004577

Hashimzade, N., Huang, Z., & Myles, G. D. (2010). Tax fraud by firms and optimal auditing. *International Review of Law and Economics*, *30*(1), 10-17. https://doi.org/10.1016/j.irle.2009.08.002

Hasija, S., Pinker, E. J., & Shumsky, R. A. (2008). Call Center Outsourcing Contracts Under Information Asymmetry. *Management Science*, *54*(4), 793-807. https://doi.org/10.1287/mnsc.1070.0804

Heckman, J. J. (1997). Instrumental Variables: A Study of Implicit Behavioral Assumptions Used in Making Program Evaluations. *The Journal of Human Resources*, *32*(3), 441-462. https://doi.org/10.2307/146178

Heckman, J. J. (2001). Micro Data, Heterogeneity, and the Evaluation of Public Policy: Nobel Lecture. *Journal of Political Economy*, *109*(4), 673-748. https://doi.org/10.1086/322086

Heckman, J. J., & Robb, R. (1985). Alternative methods for evaluating the impact of interventions. *Journal of Econometrics*, *30*(1-2), 239-267. https://doi.org/10.1016/0304-4076(85)90139-3

Heckman, J. J., & Smith, J. A. (1995). Assessing the Case for Social Experiments. *Journal of Economic Perspectives*, *9*(2), 85-110. https://doi.org/10.1257/jep.9.2.85

Hemming, K., Haines, T. P., Chilton, P. J., Girling, A. J., & Lilford, R. J. (2015). The stepped wedge cluster randomised trial: rationale, design, analysis, and reporting. *BMJ*, *350*(feb06 1), 1-7. https://doi.org/10.1136/bmj.h391

Holland, P. J., Cooper, B., & Hecker, R. (2015). Electronic monitoring and surveillance in the workplace: The effects on trust in management, and the moderating role of occupational type. *Personnel Review*, *44*(1), 161-175. https://doi.org/10.1108/PR-11-2013-0211

Holman, D. (2004). Employee Well-being in Call Centres. In S. Deery & N. Kinnie (Éd.), *Call Centres and Human Resource Management* (p. 223-244). London: Palgrave Macmillan UK. https://doi.org/10.1057/9780230288805_10

Holman, D., Batt, R., & Holtgrewe, U. (2007). *The Global Call Center Report: International Perspectives on Management and Employment (Executive Summary)*.

Holman, D., Chissick, C., & Totterdell, P. (2002). The effects of performance monitoring on emotional labor and well-being in call centers. *Motivation and Emotion*, *26*(1), 57-81.

Hölmstrom, B. (1979). Moral Hazard and Observability. *The Bell Journal of Economics*, *10*(1), 74-91. https://doi.org/10.2307/3003320

Hölmstrom, B., & Milgrom, P. (1991). Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design. *Journal of Law, Economics, and Organization*, *7*(0), 24-52. https://doi.org/10.1093/jleo/7.special_issue.24

Hurwicz, L., & Reiter, S. (2006). *Designing Economic Mechanisms*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511754258

Ichino, A., & Muehlheusser, G. (2008). How often should you open the door? *Journal of Economic Behavior & Organization*, *67*(3-4), 820-831. https://doi.org/10.1016/j.jebo.2008.02.004

Ichniowski, C., Shaw, K., & Prennushi, G. (1997). The Effects of Human Resource Management Practices on Productivity: A Study of Steel Finishing Lines. *The American Economic Review*, *87*(3), 291-313.

Imbens, G. W., & Wooldridge, J. M. (2009). Recent Developments in the Econometrics of Program Evaluation. *Journal of Economic Literature*, *47*(1), 5-86. https://doi.org/10.1257/jel.47.1.5

Jarvis, J., Dickerson, C., & Stockwell, L. (2013). Staff-student Partnership in Practice in Higher Education: The Impact on Learning and Teaching. *Procedia - Social and Behavioral Sciences*, *90*, 220-225. https://doi.org/10.1016/j.sbspro.2013.07.085

Jatteau, A. (2013). *Les expérimentations aléatoires en économie*. Paris: La Découverte.

Jensen, M. C., & Meckling, W. H. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics*, *3*(4), 305-360. https://doi.org/10.1016/0304-405X(76)90026-X

Jeske, D., & Santuzzi, A. M. (2015). Monitoring what and how: psychological implications of electronic performance monitoring. *New Technology, Work and Employment*, *30*(1), 62-78. https://doi.org/10.1111/ntwe.12039

Jost, P.-J. (1991). Monitoring in Principal-Agent Relationships. *Journal of Institutional and Theoretical Economics*, *147*(3), 517-538.

Joubert, N. (2009). Processus de détection et évaluation de la fraude sociale. *Revue économique*, *60*(5), 1235-1256. https://doi.org/10.3917/reco.605.1235

Kagel, J. H., & Roth, A. E. (Éd.). (1995). *The handbook of experimental economics*. (Vol. 1). Princeton, NJ: Princeton Univ. Press.

Kalai, E., & Smorodinsky, M. (1975). Other Solutions to Nash's Bargaining Problem. *Econometrica*, *43*(3), 513-518. https://doi.org/10.2307/1914280

Kessler, J. B., & Leider, S. (2016). Procedural Fairness and the Cost of Control. *Journal of Law, Economics, and Organization*, *32*(4), 685-718. https://doi.org/10.1093/jleo/eww009

Kim, S. K. (1995). Efficiency of an Information System in an Agency Model. *Econometrica*, *63*(1), 89-102. https://doi.org/10.2307/2951698

Kinnie, N., Purcell, J., & Adams, M. (2008). Explaining Employees' Experience of Work in Outsourced Call Centres: The Influence of Clients, Owners and Temporary Work Agencies. *Journal of Industrial Relations*, *50*(2), 209-227. https://doi.org/10.1177/0022185607087898

Kizza, J. M., & Ssanyu, J. (2005). Workplace surveillance. In J. Weckert (Éd.), *Electronic Monitoring in the Workplace: Controversies and Solutions* (p. 1-18). Hershey, PA: Idea Group Publishing.

Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, *119*(2), 254-284. https://doi.org/10.1037/0033-2909.119.2.254

Knight, F. H. (1921). *Risk, uncertainty and profit*. New York: Hart, Schaffner and Marx.

Kolm, S.-C. (1973). A note on optimum tax evasion. *Journal of Public Economics*, *2*(3), 265-270. https://doi.org/10.1016/0047-2727(73)90018-2

Koole, G. (2008). Introduction to the Special Issue on Call Center Management. *Management Science*, *54*(2), 237-237. https://doi.org/10.1287/mnsc.1080.0855

Kosonen, T., & Ropponen, O. (2015). The role of information in tax compliance: Evidence from a natural field experiment. *Economics Letters*, *129*, 18-21. https://doi.org/10.1016/j.econlet.2015.01.026

Kramarz, F., Angrist, J. D., Blau, D. M., Falk, A., Robin, J.-M., & Taber, C. R. (2006). How to do empirical economics. *Investigaciones Economicas*, *30*(2), 179-206.

Kramarz, F., & Michaud, M.-L. (2010). The shape of hiring and separation costs in France. *Labour Economics*, *17*(1), 27-37. https://doi.org/10.1016/j.labeco.2009.07.005

Labrousse, A. (2010). Nouvelle économie du développement et essais cliniques randomisés : une mise en perspective d'un outil de preuve et de gouvernement. *Revue de la régulation*, (7). https://doi.org/10.4000/regulation.7818

Laffont, J.-J., & Martimort, D. (2002). *The theory of incentives: the principal-agent model*. Princeton, N.J.: Princeton University Press. Consulté à l'adresse http://public.eblib.com/choice/publicfullrecord.aspx?p=485766

LaLonde, R. J. (1986). Evaluating the Econometric Evaluations of Training Programs with Experimental Data. *The American Economic Review*, *76*(4), 604-620.

Lambert-Le Mener, M. (2012). *La performance académique des étudiants en première année universitaire : influence des capacités cognitives et de la motivation*. Université de Bourgogne.

Lanciano-Morandat, C., Nohara, H., & Tchobanian, R. (2005). *French Call Centre Industry Report 2004*. Consulté à l'adresse https://halshs.archives-ouvertes.fr/halshs-00430100

Le Gall, R. (2015). Le contrôle des demandeurs d'emplois. *Les Cahiers Français*, (385), 80-86.

Leamer, E. E. (1983). Let's Take the Con Out of Econometrics. *The American Economic Review*, *73*(1), 31-43.

Leibenstein, H. (1980). *Beyond economic man: a new foundation for microeconomics*. Cambridge: Harvard Univ. Press.

Leslie, L. J., & Gorman, P. C. (2017). Collaborative design of assessment criteria to improve undergraduate student engagement and performance. *European Journal of Engineering Education*, *42*(3), 286-301. https://doi.org/10.1080/03043797.2016.1158791

Levine, M. F., Taylor, J. C., & Davis, L. E. (1984). Defining Quality of Working Life. *Human Relations*, *37*(1), 81-104. https://doi.org/10.1177/001872678403700105

Levinthal, D. (1988). A survey of agency models of organizations. *Journal of Economic Behavior & Organization*, *9*(2), 153-185. https://doi.org/10.1016/0167-2681(88)90071-6

Levitt, S. D., & List, J. A. (2009). Field experiments in economics: The past, the present, and the future. *European Economic Review*, *53*(1), 1-18. https://doi.org/10.1016/j.euroecorev.2008.12.001

L'Horty, Y., & Petit, P. (2011). Évaluation aléatoire et expérimentations sociales. *Revue française d'économie*, *XXVI*(1), 13-48. https://doi.org/10.3917/rfe.111.0013

List, J. A., & Reiley, D. (2010). Field Experiments. In S. N. Durlauf & L. E. Blume (Éd.), *Microeconometrics* (p. 53-58). London: Palgrave Macmillan UK. https://doi.org/10.1057/9780230280816_9

Livian, Y.-F., & Baret, C. (2002). Le contrôle de la productivité dans les activités de services. *Travail et emploi*, (91), 71-78.

Locke, E. A., & Latham, G. P. (2002). Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. *American Psychologist*, *57*(9), 705-717. https://doi.org/10.1037//0003-066X.57.9.705

Lourel, M. (2006). Rapports au travail, contrôle et santé dans les centres de gestion de la relation–client. *Psychologie du Travail et des Organisations*, *12*(1), 39-51. https://doi.org/10.1016/j.pto.2005.07.007

Maetz, I. (2017). *Parcours et réussite en licence et en Paces : les résultats de la session 2016* (Note Flash du SIES No. 11-18). MESRI - SIES.

Manchester, J. (1997). Evasion in social security: an international perspective. *Proceedings. Annual Conference on Taxation and Minutes of the Annual Meeting of the National Tax Association*, *90*, 171-177.

March, J. G., Simon, H. A., & Guetzkow, H. S. (1993). *Organizations* (2nd ed). Cambridge: Blackwell.

Marchese, C. (2005). Taxation, Black Markets, and Other Unintended Consequences. In J. G. Backhaus & R. E. Wagner (Éd.), *Handbook of Public Finance* (p. 237-275). Boston: Kluwer Academic Publishers. https://doi.org/10.1007/1-4020-7864-1_10

Mares, I. (2001). Firms and the Welfare State: When, Why, and How Does Social Policy Matter to Employers? In D. Soskice & P. A. Hall (Éd.), *Varieties of Capitalism* (p. 182-212). Oxford University Press. https://doi.org/10.1093/0199247757.001.0001

Mares, I. (2003). The Sources of Business Interest in Social Insurance: Sectoral versus National Differences. *World Politics*, *55*(02), 229-258. https://doi.org/10.1353/wp.2003.0012

Marrelli, M., & Martina, R. (1988). Tax evasion and strategic behaviour of the firms. *Journal of Public Economics*, *37*(1), 55-69. https://doi.org/10.1016/0047-2727(88)90004-7

Marx, G. T. (1999). Measuring everything that moves: the new surveillance at work. In I. H. Simpson, R. L. Simpson, & R. Hodson (Éd.), *Deviance in the workplace* (p. 165-189). Stamford, Conn: Jai Press.

Mascagni, G. (2018). From the lab to the field: a review of tax experiments. *Journal of Economic Surveys*, *32*(2), 273-301. https://doi.org/10.1111/joes.12201

Mayo, E. (2003). *The Human Problems of an Industrial Civilization*. London: Routledge.

McGillivray, W. (2001). Contribution evasion: Implications for social security pension schemes. *International Social Security Review*, *54*(4), 3-22. https://doi.org/10.1111/1468-246X.t01-1-00102

McKeachie, W. J. (1962). Procedures and Techniques of Teaching: A Survey of Experimental Studies. In N. Sanford (Éd.), *The American college: A psychological and social interpretation of the higher learning*. (p. 312-364). Hoboken: John Wiley & Sons Inc. https://doi.org/10.1037/11181-008

McMillan, J. H. (Éd.). (2013). *Handbook of research on classroom assessment*. Los Angeles: Sage Publications.

Ménard, C. (2012). *L'économie des organisations* (3. éd.). Paris: La Découverte.

Michaelis, J., & Schwanebeck, B. (2016). Examination rules and student effort. *Economics Letters*, *145*, 65-68. https://doi.org/10.1016/j.econlet.2016.05.019

Michaut, C. (2012). Réussite, échec et abandon des études dans l'enseignement supérieur français : quarante ans de recherche. In C. Michaut & M. Romainville (Éd.), *Réussite, échec et abandon dans l'enseignement supérieur* (De Boeck Supérieur, p. 53-68).

Millet, M. (2012). L'« échec » des étudiants de premiers cycles dans l'enseignement supérieur en France. Retours sur une notion ambiguë et descriptions empiriques. In C. Michaut & M. Romainville (Éd.), *Réussite, échec et abandon dans l'enseignement supérieur* (De Boeck Supérieur, p. 69-88).

Mintzberg, H. (1973). *The nature of managerial work*. New York, NY: Harper & Row.

Mirrlees, J. A. (1999). The Theory of Moral Hazard and Unobservable Behaviour: Part I. *Review of Economic Studies*, *66*(1), 3-21. https://doi.org/10.1111/1467-937X.00075

Monnier, E. (1992). *Évaluations de l'action des pouvoirs publics* (2. éd. rev. et augm). Paris: Économica.

Morlaix, S., & Lambert-Le Mener, M. (2015). La motivation des étudiants à l'entrée à l'université : quels effets directs ou indirects sur la réussite? *Recherches en Éducation*, (22), 152-167.

Morlaix, S., & Perret, C. (2013). L'évaluation du Plan Réussite en Licence : quelles actions pour quels effets ? *Recherches en Éducation*, (15), 137-150.

Mossin, J. (1968). Optimal Multiperiod Portfolio Policies. *The Journal of Business*, *41*(2), 215-229.

Nagin, D. S., Rebitzer, J. B., Sanders, S., & Taylor, L. J. (2002). Monitoring, Motivation, and Management: The Determinants of Opportunistic Behavior in a Field Experiment. *American Economic Review*, *92*(4), 850-873. https://doi.org/10.1257/00028280260344498

Nash, J. F. (1950). The Bargaining Problem. *Econometrica*, *18*(2), 155-162. https://doi.org/10.2307/1907266

Nebeker, D. M., & Tatum, B. C. (1993). The Effects of Computer Monitoring, Standards, and Rewards on Work Performance, Job Satisfaction, and Stress1. *Journal of Applied Social Psychology*, *23*(7), 508-536. https://doi.org/10.1111/j.1559-1816.1993.tb01101.x

Nyland, C., Smyth, R., & Zhu, C. J. (2006). What Determines the Extent to which Employers will Comply with their Social Security Obligations? Evidence from Chinese Firm-level Data. *Social Policy and Administration*, *40*(2), 196-214. https://doi.org/10.1111/j.1467-9515.2006.00484.x

OECD (Éd.). (2013). *Synergies for better learning: an international perspective on evaluation and assessment*. Paris: OECD.

Office Central de Lutte contre le travail illégal. (2009). *Bilan 2009*.

Palier, B. (2005). *Gouverner la sécurité sociale : les réformes du système français de protection sociale depuis 1945*. Paris: Presses universitaires de France.

Parent, G. (2014). Le contrôle de la recherche d'emploi et les sanctions. *Travail et emploi*, (139), 91-104. https://doi.org/10.4000/travailemploi.6438

Pascual, A., & Guéguen, N. (2002). La technique du « vous êtes libre de... »: Induction d'un sentiment de liberté et soumission à une requête ou le paradoxe d'une liberté manipulatrice. *Revue Internationale de Psychologie Sociale*, *15*(1), 51-80.

Pelletier, L. G., Séguin-Lévesque, C., & Legault, L. (2002). Pressure from above and pressure from below as determinants of teachers' motivation and teaching behaviors. *Journal of Educational Psychology*, *94*(1), 186-196. https://doi.org/10.1037//0022-0663.94.1.186

Perkins, D. (2013). Electronic Performance Monitoring in Call Centers: An Ethical Decision Model. *Electronic Journal of Business Ethics and Organization Studies*, *18*(1), 4-14.

Perraudin, C., Thèvenot, N., & Valentin, J. (2013). Sous-traitance et évitement de la relation d'emploi : les comportements de substitution des entreprises industrielles en France entre 1984 et 2003. *Revue internationale du Travail*, *152*(3-4), 571-597. https://doi.org/10.1111/j.1564-9121.2013.00195.x

Pfeffer, J. (1995). Incentives in organizations: The importance of social relations. In O. E. Williamson (Éd.), *Organization theory: from Chester Barnard to the present and beyond* (p. 72-97). New York: Oxford University Press.

Pickhardt, M., & Prinz, A. (2014). Behavioral dynamics of tax evasion – A survey. *Journal of Economic Psychology*, *40*, 1-19. https://doi.org/10.1016/j.joep.2013.08.006

Pitchford, R. (1998). Moral hazard and limited liability: The real effects of contract bargaining. *Economics Letters*, *61*(2), 251-259. https://doi.org/10.1016/S0165-1765(98)00141-4

Pomeranz, D. (2015). No Taxation without Information: Deterrence and Self-Enforcement in the Value Added Tax. *American Economic Review*, *105*(8), 2539-2569. https://doi.org/10.1257/aer.20130393

Pratt, J. W., & Zeckhauser, R. (Éd.). (1985). *Principals and agents: the structure of business*. Boston, Mass: Harvard Business School Press.

Reed, M. I. (2001). Organization, Trust and Control: A Realist Analysis. *Organization Studies*, *22*(2), 201-228. https://doi.org/10.1177/0170840601222002

Reeve, J., Vansteenkiste, M., Assor, A., Ahmad, I., Cheon, S. H., Jang, H., … Wang, C. K. J. (2014). The beliefs that underlie autonomy-supportive and controlling teaching: A multinational investigation. *Motivation and Emotion*, *38*(1), 93-110. https://doi.org/10.1007/s11031-013-9367-0

Reinganum, J. F., & Wilde, L. L. (1985). Income tax compliance in a principal-agent framework. *Journal of Public Economics*, *26*(1), 1-18. https://doi.org/10.1016/0047-2727(85)90035-0

Ren, Z. J., & Zhou, Y.-P. (2008). Call Center Outsourcing: Coordinating Staffing Level and Service Quality. *Management Science*, *54*(2), 369-383. https://doi.org/10.1287/mnsc.1070.0820

Rinne, R., & Koivula, J. (2005). The Changing Place of the University and a Clash of Values. In OECD, *Higher Education Management and Policy* (Vol. 17, p. 91-123). OECD Publishing. https://doi.org/10.1787/hemp-v17-art20-en

Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, Schools, and Academic Achievement. *Econometrica*, *73*(2), 417-458. https://doi.org/10.1111/j.1468-0262.2005.00584.x

Rosanvallon, P. (1995). *La nouvelle question sociale repenser l'État-providence*. Paris: Editions du Seuil. Consulté à l'adresse http://banq.pretnumerique.ca/accueil/isbn/9782021287769

Ross, S. A. (1973). The Economic Theory of Agency: The Principal's Problem. *The American Economic Review*, *63*(2), 134-139.

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, *66*(5), 688-701. https://doi.org/10.1037/h0037350

Russell, B. (2008). Call centres: A decade of research. *International Journal of Management Reviews*, *10*(3), 195-219. https://doi.org/10.1111/j.1468-2370.2008.00241.x

Rust, C., Price, M., & O'Donovan, B. (2003). Improving Students' Learning by Developing their Understanding of Assessment Criteria and Processes. *Assessment & Evaluation in Higher Education*, *28*(2), 147-164. https://doi.org/10.1080/02602930301671

Rutström, E. E. (1998). Home-grown values and incentive compatible auction design. *International Journal of Game Theory*, *27*(3), 427-441. https://doi.org/10.1007/s001820050082

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, *55*(1), 68-78. https://doi.org/10.1037//0003-066X.55.1.68

Sambell, K., McDowell, L., & Montgomery, C. (2013). *Assessment for learning in higher education*. Abingdon, Oxon; New York: Routledge. Consulté à l'adresse http://public.eblib.com/choice/publicfullrecord.aspx?p=1020266

Sandmo, A. (2005). The Theory of Tax Evasion: A Retrospective View. *National Tax Journal*, *58*(4), 643-663.

Sandmo, A. (2012). An evasive topic: theorizing about the hidden economy. *International Tax and Public Finance*, *19*(1), 5-24. https://doi.org/10.1007/s10797-011-9185-9

Schnedler, W., & Vadovic, R. (2011). Legitimacy of Control. *Journal of Economics & Management Strategy*, *20*(4), 985-1009. https://doi.org/10.1111/j.1530-9134.2011.00315.x

Scriven, M. (1967). The methodology of evaluation. In R. Tyler W., R. Gagné, & M. Scriven (Éd.), *Perspectives of currriculum evaluation* (Rand McNally, p. 39-83). Chicago.

Sewell, G., & Barker, J. R. (2006). Coercion Versus Care: Using Irony to Make Sense of Organizational Surveillance. *Academy of Management Review*, *31*(4), 934-961. https://doi.org/10.5465/amr.2006.22527466

Sewell, G., Barker, J. R., & Nyberg, D. (2012). Working under intensive surveillance: When does' measuring everything that moves' become intolerable? *Human Relations*, *65*(2), 189-215. https://doi.org/10.1177/0018726711428958

Shapiro, C., & Varian, H. R. (1999). *Information rules: a strategic guide to the network economy*. Boston, Mass: Harvard Business School Press.

Shaw, K. (2009). Insider econometrics: A roadmap with stops along the way. *Labour Economics*, *16*(6), 607-617. https://doi.org/10.1016/j.labeco.2009.09.001

Simon, H. A. (1959). Theories of Decision-Making in Economics and Behavioral Science. *The American Economic Review*, *49*(3), 253-283.

Simon, H. A. (1991). Bounded Rationality and Organizational Learning. *Organization Science*, *2*(1), 125-134. https://doi.org/10.1287/orsc.2.1.125

Slemrod, J. (1985). An Empirical Test for Tax Evasion. *The Review of Economics and Statistics*, *67*(2), 232-238. https://doi.org/10.2307/1924722

Slemrod, J. (2007). Cheating Ourselves: The Economics of Tax Evasion. *The Journal of Economic Perspectives*, *21*(1), 25-48.

Slemrod, J., & Weber, C. (2012). Evidence of the invisible: toward a credibility revolution in the empirical analysis of tax evasion and the informal economy. *International Tax and Public Finance*, *19*(1), 25-53. https://doi.org/10.1007/s10797-011-9181-0

Slemrod, J., & Yitzhaki, S. (2002). Chapter 22: Tax Avoidance, Evasion, and Administration. In *Handbook of Public Economics* (Vol. 3, p. 1423-1470). Elsevier. https://doi.org/10.1016/S1573-4420(02)80026-X

Sliwka, D. (2007). Trust as a Signal of a Social Norm and the Hidden Costs of Incentive Schemes. *American Economic Review*, *97*(3), 999-1012. https://doi.org/10.1257/aer.97.3.999

Smith, M. J., Carayon, P., Sanders, K. J., Lim, S.-Y., & LeGrande, D. (1992). Employee stress and health complaints in jobs with and without electronic performance monitoring. *Applied Ergonomics*, *23*(1), 17-27. https://doi.org/10.1016/0003-6870(92)90006-H

Sneyers, E., & De Witte, K. (2018). Interventions in higher education and their effect on student success: a meta-analysis. *Educational Review*, *70*(2), 208-228. https://doi.org/10.1080/00131911.2017.1300874

Spence, M. (1973). Job Market Signaling. *The Quarterly Journal of Economics*, *87*(3), 355-374. https://doi.org/10.2307/1882010

Srinivasan, T. N. (1973). Tax evasion: A model. *Journal of Public Economics*, *2*(4), 339-346. https://doi.org/10.1016/0047-2727(73)90024-8

Stahl, B. C., Prior, M., Wilford, M., & Collins, D. (2005). Electronic Monitoring in the Workplace: If People Don't Care, Then What is the Relevance? In J. Weckert (Éd.), *Electronic Monitoring in the Workplace: Controversies and Solutions* (p. 50-78). IGI Global. https://doi.org/10.4018/978-1-59140-456-9

Stanton, Jeffrey M. (2000). Reactions to Employee Performance Monitoring: Framework, Review, and Research Directions. *Human Performance*, *13*(1), 85-113. https://doi.org/10.1207/S15327043HUP1301_4

Stanton, Jeffrey M, & Julian, A. L. (2002). The impact of electronic monitoring on quality and quantity of performance. *Computers in Human Behavior*, *18*(1), 85-101. https://doi.org/10.1016/S0747-5632(01)00029-2

Stanton, Jeffrey M., & Sarkar-Barney, S. T. M. (2003). A Detailed Analysis of Task Performance With and Without Computer Monitoring. *International Journal of Human-Computer Interaction*, *16*(2), 345-366. https://doi.org/10.1207/S15327590IJHC1602_11

Stanton, J.M, & Weiss, E. (2000). Electronic monitoring in their own words: an exploratory study of employees' experiences with new types of surveillance. *Computers in Human Behavior*, *16*(4), 423-440. https://doi.org/10.1016/S0747-5632(00)00018-2

Stefani, L. A. J. (1998). Assessment in Partnership with Learners. *Assessment & Evaluation in Higher Education*, *23*(4), 339-350. https://doi.org/10.1080/0260293980230402

Stiglitz, J. E. (2002). Information and the Change in the Paradigm in Economics. *American Economic Review*, *92*(3), 460-501. https://doi.org/10.1257/000282802760136363

Taylor, P., & Bain, P. (1999). « An assembly line in the head »: work and employee relations in the call centre. *Industrial Relations Journal*, *30*(2), 101-117. https://doi.org/10.1111/1468-2338.00113

Taylor, P., Mulvey, G., Hyman, J., & Bain, P. (2002). Work Organization, Control and the Experience of Work in Call Centres. *Work, Employment and Society*, *16*(1), 133-150. https://doi.org/10.1177/09500170222119281

Théophile, N. (2017). *Réussite en licence : la moitié des entrants valide une licence* (Note Ofipe No. 33). Observatoire des formations, des Insertions Professionnelles, Évaluations.

Thépaut, Y. (2006). Le concept d'information dans l'analyse économique contemporaine. *Hermès, La Revue*, *44*(1), 161-168.

Tian, D. (2011). *Rapport d'information sur la lutte contre la fraude sociale*. Assemblée Nationale.

Tomczak, D. L., Lanzo, L. A., & Aguinis, H. (2018). Evidence-based recommendations for employee performance monitoring. *Business Horizons*, *61*(2), 251-259. https://doi.org/10.1016/j.bushor.2017.11.006

Torgler, B. (2002). Speaking to Theorists and Searching for Facts: Tax Morale and Tax Compliance in Experiments. *Journal of Economic Surveys*, *16*(5), 657-683. https://doi.org/10.1111/1467-6419.00185

Torgler, B. (2008). What Do We Know about Tax Fraud? An Overview of Recent Developments. *Social Research*, *75*(4,), 1239-1270.

van Ackere, A. (1993). The principal/agent paradigm: Its relevance to various functional fields. *European Journal of Operational Research*, *70*(1), 83-103. https://doi.org/10.1016/0377-2217(93)90234-E

Viet, V. (s. d.). Protection Sociale. In *Encyclopædia Universalis*. Consulté à l'adresse https://www-universalis--edu-com.fennec.u-pem.fr/encyclopedie/protection-sociale/

Vincent, A. L.-A. (1965). De la mesure de la productivité aux problèmes généraux de gestion économique. *Revue économique*, *16*(6), 879-924. https://doi.org/10.3406/reco.1965.407685

Vlassis, M., & Mamakis, S. (2014). *Undeclared labour in Cournot Oligopoly*.

von Siemens, F. A. (2013). Intention-based reciprocity and the hidden costs of control. *Journal of Economic Behavior & Organization*, *92*, 55-65. https://doi.org/10.1016/j.jebo.2013.04.017

Wang, L. F. S., & Conant, J. L. (1988). Corporate tax evasion and output decisions of the uncertain monopolist. *National Tax Journal*, *41*(4), 579-581.

Wellen, J., Martin, A., & Hanson, D. (2009). The impact of electronic surveillance and workplace empowerment on work attitudes and behaviour. Présenté à 8th Industrial and Organisational Psychology Conference (IOP).

West, J. P., & Bowman, J. S. (2016). Electronic Surveillance at Work: An Ethical Analysis. *Administration & Society*, *48*(5), 628-651. https://doi.org/10.1177/0095399714556502

Wickham, J., & Collins, G. (2004). The call centre: a nursery for new forms of work organisation? *The Service Industries Journal*, *24*(1), 1-18. https://doi.org/10.1080/02642060412331301102

Williamson, O. E. (1967). Hierarchical Control and Optimum Firm Size. *Journal of Political Economy*, *75*(2), 123-138. https://doi.org/10.1086/259258

Williamson, O. E. (1975). *Markets and hierarchies, analysis and antitrust implications: a study in the economics of internal organization*. New York: Free Press.

Wu, L., Waber, B. N., Aral, S., Brynjolfsson, E., & Pentland, A. (2008). *Mining Face-to-Face Interaction Networks using Sociometric Badges: Predicting Productivity in an IT Configuration Task* (SSRN Scholarly Paper No. ID 1130251). Rochester, NY: Social Science Research Network. Consulté à l'adresse https://papers.ssrn.com/abstract=1130251

Yaniv, G. (1988). Withholding and non-withheld tax evasion. *Journal of Public Economics*, *35*(2), 183-204. https://doi.org/10.1016/0047-2727(88)90053-9

Yaniv, G. (1995). A note on the tax-evading firm. *National Tax Journal*, *48*(1), 113-120.

Yitzhaki, S. (1974). Income tax evasion: A theoretical analysis. *Journal of Public Economics*, *3*(2), 201-202. https://doi.org/10.1016/0047-2727(74)90037-1

Zajonc, R. B. (1965). Social Facilitation. *Science*, *149*(3681), 269-274.

Zand, D. E. (1972). Trust and Managerial Problem Solving. *Administrative Science Quarterly*, *17*(2), 229-239. https://doi.org/10.2307/2393957

Zaugg, V., Savoldelli, V., Sabatier, B., & Durieux, P. (2014). Améliorer les pratiques et l'organisation des soins : méthodologie des études d'interventions. *Santé Publique*, *26*(4), 519-529. https://doi.org/10.3917/spub.144.0519

Ziegelmeyer, A., Schmelz, K., & Ploner, M. (2012). Hidden costs of control: four repetitions and an extension. *Experimental Economics*, *15*(2), 323-340. https://doi.org/10.1007/s10683-011-9302-8

Zubrickas, R. (2015). Optimal grading. *International Economic Review*, *56*(3), 751-776. https://doi.org/10.1111/iere.12121

L'ÉCONOMIE DES DISPOSITIFS DE VÉRIFICATION DE L'INFORMATION : UNE APPROCHE EXPÉRIMENTALE

L'économie des contrats complets prédit qu'au sein d'une relation d'agence d'une organisation productive, en présence d'aléa moral, un dispositif de vérification de l'information permet de répondre à la fois à un problème de coopération entre les individus et à un problème de coordination des activités. Cependant, au lieu de discipliner des comportements opportunistes, ce dispositif peut engendrer des coûts cachés et réduire la motivation intrinsèque des agents à réaliser une activité qui leur a été attribuée. Sous certaines conditions, il génère une réduction de l'activité, et une perte en termes d'allocation ce qui nuit à l'efficacité.

Dans cette thèse de doctorat, nous avons conduit trois expérimentations contrôlées de terrain avec assignation aléatoire qui visaient à modifier les configurations du dispositif de vérification de l'information afin de résoudre un problème organisationnel propre à trois relations d'agence particulières.

Dans notre premier chapitre, nous avons testé l'effet de la variation de la quantité d'informations détenues par les cotisants sur le dispositif de vérification de la déclaration sociale grâce à des messages ciblés contenant des explications sur le pouvoir de contrôle de l'Agence centrale des organismes de sécurité sociale (Acoss) afin de réduire la fraude sociale.

Dans notre deuxième chapitre, nous avons testé la réduction de l'intensité de la surveillance électronique de la performance des conseillers d'un centre d'appels sous-traitants afin d'améliorer leur qualité de vie au travail.

Enfin, dans notre troisième chapitre, nous avons testé l'effet de la négociation contractuelle du dispositif d'évaluation de la participation des étudiants de licence pendant les travaux dirigés afin d'améliorer leur réussite à l'université.

THE ECONOMICS OF INFORMATION CHECK DEVICES: AN EXPERIMENTAL APPROACH

Economics of complete contracts foresees that within an agency relationship of a productive organization, in the presence of moral hazard, an information check device can address both a problem of cooperation between individuals and a problem of coordination of activities. However, instead of disciplining opportunistic behaviours, this device can generate hidden costs and reduce the intrinsic motivation of agents to perform an activity that has been assigned to them. Under certain conditions, it generates a reduction of the outcome, and a loss in terms of allocation, which is detrimental to efficiency.

In this Ph.D. thesis, we conducted three randomized controlled field trials that aimed at modifying the configurations of the information check device to solve an organizational problem specific to three specific agency relationships.

In our first chapter, we tested the effect of varying the amount of information held by contributors on the social reporting verification device through targeted messages containing explanations of the control power of the Agence centrale des organisations de sécurité sociale (Acoss) in order to reduce social fraud.

In our second chapter, we tested the reduction in the intensity of the electronic monitoring of the performance of advisors of an outsourced call centre in order to improve their quality of life at work.

Finally, in our third chapter, we tested the effect of the contractual negotiation of the device which evaluates the participation of undergraduate students during the tutorials in order to improve their success at the university.

UNIVERSITÉ PARIS-EST
ÉCOLE DOCTORALE ORGANISATIONS, MARCHÉS, INSTITUTIONS (ED 530)