Essays on the economics of social identity, social preferences and social image

Rémi Suchon

Numéro National de Thèse : 2018LYSEN080

THESE de DOCTORAT DE L'UNIVERSITE DE LYON
opérée par
l'Ecole Normale Supérieure de Lyon

Ecole Doctorale N° 486
Sciences Economiques et de Gestion

Spécialité de doctorat : Economie

Soutenue publiquement le 14/12/2018, par :

Rémi SUCHON

Essays on the Economics of Social Identity, Social Preferences and Social Image

Essais en économie de l'identité sociale, des préférences sociales et de l'image sociale

---

Devant le jury composé de :

CORGNET, Brice - Professeur, EM Lyon	Examinateur
HAMMAN, John - Associate Professor, Florida State University, USA	Examinateur
JACQUEMET, Nicolas - Professeur, Université Paris I et Paris School of Economics	Rapporteur
TERGIMAN, Chloé – Assistant Professor, The Pennsylvania State University,USA	Examinatrice
VILLEVAL, Marie Claire – Directrice de recherche CNRS Université de Lyon	Directrice de thèse
WILLINGER, Marc – Professeur, Université Montpellier I -	Rapporteur

Ecole Normale Supérieure de Lyon is not going to give any approbation or disapprobation about the thoughts expressed in this dissertation. They are only the author's ones and need to be considered such as.

Remerciement

Ces quatres dernières années, consacrées en grande partie à l'apprentissage du métier de chercheur, ont été par moment très éprouvantes mais toujours riches d'enseignements. Je ne considère pas le manuscrit qui suit comme l'aboutissement de cet apprentissage, mais plutôt comme un étape de ma formation intellectuelle qui ne fait que commencer.

Cette formation intellectuelle a été fortement marquée et infléchie par Marie Claire Villeval qui m'a encadré et soutenu au long de cette thèse. A ses côtés, j'ai appris l'exigence nécessaire à la production de recherche scientifique, la rigueur et la méticulosité, mais également que le moteur essentiel de la démarche doit être la passion et la curiosité. Marie Claire: merci pour tout cela.

Je remercie Brice Corgnet, John Hamman et Chloé Tergiman d'avoir accepté de faire partie de mon jury de thèse. Je remercie Nicolas Jaquemet et Marc Willinger d'avoir accepté le rôle de rapporteur. Je suis très heureux de présenter mon travail devant vous.

Je remercie Dan Houser de m'avoir acceuilli à l'Interdisciplinary Center for Economic Science pendant 6 mois. Ce séjour aura été fertile puisque l'idée du second chapitre de la présente thèse a germé dans un bureau du Vernon Smith Hall.

Pour avoir apporté un soutien financier indispensable à la réalisation de mes expériences, ainsi qu'à mon séjour aux Etats-unis, je remercie le Labex Cortex.

J'ai eu la chance de réaliser cette thèse au GATE. Le laboratoire fournit des conditions idéales aux doctorants, à la fois sur le plan matériel et sur le plan humain. Merci à sa directrice, Sonia Paty, qui se préoccupe de tout cela. Merci aux membres de l'équipe technique et administrative, toujours prêts à aider et surtout toujours sympathiques. Merci aux enseignants chercheurs du laboratoire, en particuliers aux expérimentalistes dont l'expertise m'a souvent inspiré. Merci à Quentin d'avoir désamorcé la tension d'un bon vieux "c'est normal" quand une force obscure paralysait les ordinateurs de la salle d'expérience au moment où les participants arrivaient.

Travailler au GATE pendant de longues (parfois très longues) journées aurait été moins agréable en l'absence de camarades idoines. (Sans ordre de préférence ;)) Vincent, Julien, Charlotte, Tidiane, Thomas, Valentin, Yohann, Liza, Tatiana, Morgan, Clément, Maxime, Chrisa, Marius, Jocelyn, Wilfried, Alice, Claire, Benjamin, Tiruo, Chloé: Merci ! c'était bien chouette. Je serais bientôt disposé à remettre mon titre de vice-champion de ping-pong du GATE en jeu.

En dehors du cercle académique, je remercie ma famille pour n'avoir pas trop cherché à comprendre: je vous expliquerai tout bientôt. Merci à mes amis également. Les petites B. ensemble sont toujours aussi délicieuses.

Z, tu es plus pour moi que ce que quelques mots ici peuvent exprimer. Mais puisque tu l'as subtilement suggéré et que je ne peux rien te refuser, merci à toi. Ta force m'inspire, ton éclat m'illumine.

Résumé de la thèse

Cette thèse étudie trois déterminants sociaux des décisions économiques : l'identité sociale, l'image sociale et les préférences sociales. Les individus prennent des décisions dans des contextes sociaux et l'économie, en tant que science sociale des choix, s'attache de plus en plus à comprendre comment les facteurs sociaux influencent les décisions économiques.

Dans l'œuvre fondatrice d'Adam Smith (1759) , l'altruisme, l'équité ou l'estime sociale étaient déjà identifiés parmi les déterminants des décisions économiques (Ashraf et al., 2005) : dans certaines circonstances, les comportements individuels ne sont pas régis par des passions, qui conduisent naturellement à l'égoïsme, mais par le "spectateur impartial" qui cherche à se conformer aux règles sociales. Vernon Smith (2012, 2013) fait référence aux idées contenues dans l'œuvre fondatrice d'Adam Smith par le terme "Humanomics". Beaucoup de comportements économiques restent incompréhensibles si nous ignorons que les individus cherchent à être dignes de l'estime d'autrui et évitent les comportements socialement répréhensibles, et s'ils ne sont pas analysés à la lumière des règles sociales émergentes qui définissent quels comportements sont répréhensibles et quels comportements sont dignes d'estime.

Cependant, l'analyse économique a quelque peu relégué ces motivations au second plan. Dans sa définition de l'Homme Economique, John Stuart Mill a décrit les motivations derrières les décisions économiques comme étant matérielles, égoïstes et

asociales. La figure de l'Homme Economique a longtemps dominé l'univers conceptuel des économistes, parce qu'elle constituait un cadre simple et puissant qui fournissait d'importantes perspectives sur une variété de questions économiques. Les motivations sociales peuvent aussi être incompatibles avec le réductionnisme, pilier important de la méthodologie de l'économie positive (Friedman, 1953), ce qui pourrait expliquer pourquoi ce type de motivations sociales a été considéré comme de second ordre dans la compréhension des choix économiques. En d'autres termes, l'Homme Economique a survécu malgré ses faiblesses car l'analyse de ses choix est suffisante pour expliquer un grand nombre de comportements économiques de manière claire et convaincante.

Outre ses faiblesses et son apparente asocialité, l'Homme Economique présente de grandes forces pour expliquer les comportements sociaux de façon parcimonieuse. Tout d'abord, la simplicité des hypothèses sur les motivations de l'Homme Economique permet une étude épurée des interactions sociales. La théorie des jeux, l'étude des interactions stratégiques, formalise pourquoi les décisions individuelles ne peuvent être comprises sans au moins une certaine prise en compte du contexte social. Un joueur qui interagit avec d'autres joueurs doit tenir compte des préférences des autres joueurs, car sa stratégie optimale dépend des stratégies des autres qui, à leurs tours, dépendent de leurs préférences.

En outre, le calcul économique censé guider les décisions de l'Homme économique peut être appliqué à des décisions (supposément) non économiques. Ainsi, l'économie de la famille de Gary Becker est une étape importante dans l'histoire de la pensée économique. Becker (1974) dote l'Homme Economique d'une certaine forme de motivation "sociale" : dans sa théorie, l'Homme Economique prend en compte le bien-être de sa progéniture dans son calcul économique. L'altruisme est alors reconnu comme une motivation légitime de l'Homme Economique, les legs s'expliquant par exemple par le souci du chef de famille pour le bien-être des héritiers. Cependant, l'altruisme chez Becker peut être considéré comme minimaliste car il est supposé influencer le

comportement dans un ensemble très restreint de décisions : les individus peuvent être altruistes dans le cadre de la famille, mais l'égoïsme reste l'hypothèse privilégiée pour expliquer les interactions sur le marché.

D'autre part, l'intégration d'éléments sociaux dans la motivation de l'homme économique a fourni de nouvelles perspectives sur des problèmes économiques classiques. Par exemple, Duesenberry (1949) a suggéré que la consommation remplit un rôle de démonstration : une consommation ostenstible fournit un statut social. Cela peut expliquer pourquoi l'épargne globale n'augmente pas comme prévu lorsque le revenu global augmente. Akerlof (1982) et Akerlof and Yellen (1988, 1990) ont importé des idées de la psychologie sociale pour expliquer la rigidité des salaires et le chômage : ils montrent que si les travailleurs sont regardant quant à l'équité de leur rémuneration, les entreprises ne peuvent pas ajuster les salaires à la baisse en cas de ralentissement économique et cela conduit à des déséquilibres sur le marché du travail. Le souci d'équité, une motivation sociale, est ainsi un antécédent convaincant du chômage. Cela démontre que les motivations sociales peuvent avoir des implications importantes y compris sur le marché.

Plus récemment, une accumulation de données issues d'expériences de laboratoire et de terrain a mis en évidence un certain nombre de résultats empiriques difficilement compatibles avec l'hypothèse de l'Homme Economique, et qui montrent de manière frappante les limites de son pouvoir prédictif. Par exemple : Pourquoi les monopoles n'utilisent-ils pas leur pouvoir de négociation pour fixer des prix élevés (Kahneman et al., 1986; Fehr et al., 1993) ? Pourquoi les individus n'exploitent-ils pas pleinement les occasions de mentir ou de tricher et laissent ainsi de l'argent sur la table (Gneezy, 2005; Fischbacher and Föllmi-Heusi, 2013) ? Pourquoi les individus donnent-ils de l'argent à des organismes de bienfaisance lorsqu'on le leur demande, mais utilisent souvent des stratégies coûteuses pour éviter d'être sollicités (Andreoni et al., 2017) ? Pourquoi les individus agissent-ils parfois de manière moins prosociale quand ils ont des motivations extrinsèques à le faire que quand ils n'en n'ont pas (Ariely et al., 2009)? Pourquoi les gens

investissent-ils des efforts coûteux pour atteindre des statuts qui ne procurent aucun gain économique (Charness et al., 2014) ? Friedman (1953) a soutenu que la qualité d'une hypothèse théorique est indépendante de son réalisme, mais est déterminée par sa capacité à prédire le comportement. Dans des décisions très simples, par exemple tricher pour de l'argent, l'hypothèse de l'homme économique n'est pas en mesure de prédire le comportement. Un examen minutieux de la pertinence d'hypothèses supplémentaires ou alternatives à propos de ce qui motive les comportements économiques est nécessaire pour améliorer la capacité de la théorie économique à comprendre les phénomènes économiques. Parmi ces hypothèses, nous nous concentrerons sur celles liées aux motivations sociales.

Dans cette thèse, nous étudions l'identité sociale, l'image sociale et les préférences sociales, trois motivations sociales d'intérêt pour l'économie car elles aident à comprendre les régularités empiriques présentées dans le paragraphe précédent. Par exemple, la volonté de préserver une bonne image sociale explique en partie pourquoi les gens n'exploitent pas pleinement les possibilités de tricherie : certaines personnes restent honnêtes pour garder l'estime des autres. La prise en compte de l'identité sociale explique pourquoi les gens sont prêts à sacrifier des ressources pour atteindre un statut élevé qui n'apporte aucun avantage économique : celà leur permet de définir *leur identitité* d'une manière positive. Les préférences sociales peuvent expliquer pourquoi les entreprises n'exploitent pas leur pouvoir de négociation : les clients qui n'aiment pas être traités injustement les boycotteraient. Les sections qui suivent passent en revue la littérature sur l'identité sociale, l'image sociale et les préférences sociales en économie en détail et présentent nos contributions à chacune de ces littératures.

Identité sociale

L'identité est la définition qu'une personne donne d'elle-même, ou en d'autres termes la réponse qu'elle apporte à la question "Qui suis-je" ? Les individus cherchent à définir et

à signaler leurs identités par leurs comportements. Par conséquent, pour bien comprendre les décisions économiques, il est important de tenir compte du sens que les individus tentent de donner à leurs comportements et de la façon dont ces comportements contribuent à construire leurs identités. Par exemple, Bénabou and Tirole (2006) démontrent qu'agir de façon prosociale est un moyen de signaler l'importance que l'on apporte à certaines vertues, ce qui permet de construire son identité. Ils présentent une implication intéressante de celà : les motivations extrinsèques peuvent réduire la propension à agir de manière prosociale, parce qu'elles brouillent le signal concernant les motivations intrinsèques de chacun. La prise en compte des motivations liées à l'identité améliore la compréhension du pouvoir et des limites des incitations.

Une partie importante de l'identité d'une personne provient de son appartenance à des groupes sociaux, ce qui constitue son identité sociale (Tajfel and Turner, 1979). L'identification aux groupes sociaux a un impact considérable sur le comportement. Un exemple frappant est la facilité avec laquelle les chercheurs peuvent générer de la discrimination fondée sur l'hostilité en laboratoire : de nombreuses expériences économiques séparent arbitrairement les participants en groupes saillants et constatent un biais intra-groupe, c'est-à-dire une tendance des participants à favoriser ceux avec qui ils partagent une identité de groupe commune (Chen and Li, 2009; Hargreaves Heap and Zizzo, 2009). Les identités de groupe induites dans le laboratoire ont également un impact sur les comportements stratégiques dans les jeux de coordination et de coopération (Charness et al., 2007; Chen and Chen, 2011). Les résultats obtenus avec des identités de groupe induites en laboratoire sont reproduits dans des études utilisant des groupes naturels : par exemple, un biais intra-groupe est constaté entre des groupes délimités par l'appartenance à une entreprise (Montmarquette et al., 2004), la religion (Hedegaard and Tyran, 2018) ou l'appartenance ethnique (Cettolin and Suetens, 2018). Il est intéressant de noter que les groupes naturels n'induisent pas nécessairement plus de biais que les groupes induits en laboratoire (Lane, 2016).

L'identité sociale a des implications économiques qui vont au-delà du simple biais intra-groupe. En renforçant l'indentification aux groupes, les chercheurs constatent que les membres ont tendance à se conformer aux stéréotypes du groupe. Par exemple, lorsque leur appartenance au groupe est rendue saillante, les Américains d'origine Asiatique deviennent plus patients (Benjamin et al., 2010), les détenus deviennent plus malhonnêtes (Cohn et al., 2015) et les religieux deviennent plus généreux (Shariff and Norenzayan, 2007).

Le chapitre 1 montre que le concept d'identité sociale offre des perspectives intéressantes sur une question économique importante. Notre question de recherche originale est d'étudier comment la mobilité sociale, définie comme le passage d'un niveau social à un niveau supérieur, influe sur la confiance en présence d'identité de groupe. Un exemple concret de mobilité ascendante est la promotion au sein des entreprises. Les promotions ont des effets sur la motivation, et l'absence de promotion peut réduire les incitations des travailleurs à fournir des efforts. Cependant, le processus de promotion peut interagir avec l'identité sociale : notre postulat est que, au sein d'un groupe social, la promotion augmente la distance sociale entre ceux qui sont promus et ceux qui ne le sont pas. D'autre part, la mobilité sociale peut réduire la distance sociale entre les promus et les membres des groupes qui ont déjà un statut élevé. Nous supposons que la mobilité ascendante pourrait avoir un impact sur la confiance par le biais de la distance sociale. L'effet de la mobilité ascendante sur la confiance est important en pratique, car la confiance entre collègues de travail est un substitut au contrôle, potentiellement coûteux, peut atténuer le problème des contrats incomplets et, en tant que tel, constitue un atout pour le succès des entreprises.

Afin de mesurer les effets de la mobilité ascendante sur la confiance, nous étudions expérimentalement un jeu de confiance (*trust game*) standard (Berg et al., 1995) dans lequel les joueurs sont identifiés par une identité de groupe naturelle et un statut. L'identité de groupe est conférée par l'école des sujets: soit l'école de commerce locale,

soit l'école d'ingénieurs locale. Les deux écoles sont sélectives et ont des identités culturelles spécifiques. Par conséquent, les sujets s'identifient fortement à leurs écoles. L'identité naturelle des sujets est renforcée dans la première partie de l'expérience au moyen d'un questionnaire sur leurs écoles respectives.

En revanche, le statut est induit en laboratoire. Nous demandons aux sujets de répondre à un questionnaire de mathématiques. La performance relative dans cette tâche détermine le rôle des sujets dans la dernière partie de l'expérience. Nous appellons "experts" les sujets qui donnent le plus grand nombre de bonnes réponses, et "agents" les autres sujets. Cela introduit une hiérarchie entre les rôles. Une caractéristique essentielle du design expérimental est que l'école prédit les performances au test de mathématiques. En conséquence, l'école prédit le statut des sujets, mais de manière imparfaite. Les sujets de l'école d'ingénieurs sont beaucoup plus susceptibles d'atteindre le statut élevé, et nous représentons la mobilité ascendante en attribuant le statut d'expert également aux sujets de l'école de commerce qui ont atteint des performances élevées. Il s'agit là d'une caractéristique importante de la mobilité ascendante, car ces sujets atteignent un statut élevé, alors que celà était peu probable ex ante compte tenu des spécificités de leur école. Il est importante de noter que, dans notre expérience, le statut est déconnecté des gains : les experts et les agents ont des rôles différents mais reçoivent les mêmes gains. Cela isole l'effet pur du statut et de l'identité sur la confiance.

Ensuite, les sujets jouent à un jeu de confiance en méthode stratégique. Ils prennent des décisions à la fois en tant que trustor et trustee, étant informés de l'école et du statut d'expert ou d'agent de leur homologue, et dans toutes les combinaisons possibles. Afin de mesurer l'impact de la mobilité ascendante sur la confiance, nous comparons les transferts d'experts de l'école de commerce à ceux des agents de l'école de commerce. Nous comparons également les transferts destinés aux experts de l'école de commerce à ceux destinés aux agents de l'école de commerce afin de mesurer l'impact de la mobilité ascendante sur la propension des autres à faire confiance et à être dignes de confiance

vis-à-vis des sujets promus.

Les résultats montrent que la mobilité ascendante a un impact spécifique sur la confiance : les sujets promus ont tendance à faire moins confiance, quelle que soit l'identité du trustee. D'autre part, les promus n'attirent pas de comportements discriminatoires : ils reçoivent les mêmes transferts que les autres sujets de l'école de commerce. Ces résultats suggèrent que l'effet de la mobilité ascendante sur la distance sociale est subtil. Les promus semblent se distancier à la fois des sujets avec qui ils partagent une identité commune et des autres, ce qui s'explique par l'expérience spécifique d'avoir le statut élevé tout en étant de l'école de commerce : les experts de l'école de commerce perçoivent probablement le statut d'expert comme plus gratifiant pour eux que pour les sujets de l'école d'ingénieurs. Par conséquent, ils se considèrent comme "singuliers" : leur statut leur donne une identité spécifique.

Le chapitre 1 apporte également une contribution plus fondamentale à la littérature sur l'identité sociale. Les résultats indiquent qu'un statut induit en laboratoire, qui n'apporte aucun gain monétaire et ne dure que le temps de l'expérience, peut défaire une identité naturelle. Cela confirme que le statut est une source d'identité puissante.

Image sociale:

Le deuxième déterminant social étudié dans cette thèse est l'image sociale, c'est-à-dire la perception des caractéristiques intrinsèques d'un individu par un tiers. L'approche économique de l'image sociale est liée au développement de l'économie de l'information. Dans toute interaction économique où l'information sur les parties prenantes est imparfaite, les perceptions sur les parties prenantes sont substituées à l'information exacte qui est indisponible. La discrimination fondée sur les croyances (Phelps, 1972; Arrow, 1973), dans laquelle l'image sociale des individus est alimentée par des stéréotypes à propos des groupes auxquels ils appartiennent, illustre clairement pourquoi l'image sociale

a des implications économiques importantes : ceux qui sont victimes de stéréotypes subissent des coûts économiques considérables.

L'image sociale entre dans la fonction d'utilité de deux façons distinctes. Tout d'abord, elle a un impact indirect sur l'utilité, par le biais de l'impact de l'image sociale sur les interactions avec les autres. Les individus prennent des dispositions coûteuses pour signaler leurs types (Riley, 2001) et pour préserver leurs réputations (Wilson, 1985). Les étudiants obtiennent des diplômes (en partie) pour être perçus comme compétents par les employeurs (Spence, 1973) ou, à l'inverse, négligent leur éducation pour démontrer leur conformité à une culture de faible réussite scolaire (Bursztyn and Jensen, 2015). Les religieux adoptent des coutumes économiquement coûteuses pour signifier la force de leur engagement vis-à-vis du culte, et bénéficier des interactions avec ses membres (Berman, 2000; Carvalho, 2013). Lorsque l'information sur les comportements passés est disponible, les individus indignes de confiance imitent les individus dignes de confiance afin de bénéficier de l'image d'invididus fiables dans les interactions ultérieures (Camerer and Weigelt, 1988; Brandts and Figueras, 2003; Huck et al., 2012).

L'image sociale entre aussi directement dans la fonction d'utilité, car beaucoup de gens se soucient de la façon dont ils sont perçus par les autres. Dans les modèles théoriques, le goût de l'estime sociale a été utilisé pour expliquer les comportements pro-sociaux (Harbaugh, 1998; Bénabou and Tirole, 2006; Andreoni and Bernheim, 2009), l'obéissance aux normes (Bernheim, 1994) ou l'honnêteté (Abeler et al., 2018; Kajackaite et al., 2018) entre autres. La littérature empirique, et en particulier les expériences variant le degré d'observabilité des décisions non stratégiques, soutient clairement l'hypothèse selon laquelle les individus attachent un poids important aux perceptions des autres, même lorsque ces perceptions ne peuvent avoir de conséquence matérielle. Hoffman et al. (1994) montrent que l'utilisation d'une procédure dans laquelle ni l'expérimentateur ni les destinataires ne peuvent identifier le dictateur réduit considérablement les transferts par rapport à une procédure dans laquelle seuls les expérimentateurs peuvent

identifier le dictateur. Andreoni and Petrie (2004) varient entre conditions l'utilisation de photographies pour identifier les joueurs dans une jeu de bien public, et constatent que les contributions sont plus importantes lorsque les images sont utilisées. Dans une tâche d'effort réel où les performances déterminent le paiement à un organisme de bienfaisance, Ariely et al. (2009) montrent que les participants ont de meilleures performances lorsqu'ils doivent révéler publiquement leurs performances aux autres participants que lorsque les performances restent privées. Fischbacher and Föllmi-Heusi (2013) demandent aux participants de reporter le résultat d'un jet de dé qu'ils observent en privé et ce report conditionne leurs gains. La distribution des reports montre qu'il existe du mensonge partiel, ce qui indique que les participants adoptent des stratégies sophistiquées pour éviter d'être perçus comme malhonnêtes.

Le chapitre 2 apporte une contribution originale à l'économie de l'image sociale. L'image sociale est généralement considérée comme un " bien privé " : les individus agissent de façon plus prosociale ou plus honnête pour être perçus positivement en tant qu'individus. La nouveauté de ce chapitre est d'étudier expérimentalement une situation dans laquelle l'image sociale a des propriétés de bien public : le comportement individuel peut se répercuter sur l'image sociale des groupes. Le jeu que nous étudions expérimentalement est à trois joueurs. Deux joueurs sont membres d'une paire, et le troisième s'appelle "l'Audience". Premièrement, les deux membres de la paire déclarent la valeur d'une variable aléatoire indépendamment et sans communication. Ils ont parfois des incitations pécuniaires à faire de fausses déclarations. Après que les deux membres de la paire ont fait leur déclaration, l'Audience observe la déclaration d'un membre prédéterminé de la paire et la valeur réelle de la variable aléatoire, de sorte qu'elle peut repérer les déclarations malhonnêtes. Ensuite, l'Audience doit décider si elle veut jouer avec le deuxième membre de la paire. La fonction de gain de l'Audience est telle qu'elle ne veut jouer que si elle estime que la probabilité d'une déclaration honnête par le second membre est suffisamment élevée. Les deux membres de la paire préfèrent que l'Audience joue, parce que quand c'est le cas, ils obtiennent une récompense supplé-

mentaire. Entre conditions, nous varions si le comportement du premier membre de la paire est informatif à propos du comportement du second membre, ce qui est important pour l'Audience qui doit former une croyance sur le second membre de la paire. Dans la condition de base, les membres de la paire ne se connaissent pas. Dans le traitement, les membres de la paire sont des amis, ce qui rend le comportement du premier membre potentiellement informatif sur le comportement du second. En effet, les décisions de deux amis sont corrélées puisque les amis ont tendance à être groupés par préférence à cause de l'homophilie et de l'influence mutuelle.

Les données montrent que l'image sociale est en effet un bien public dans les paires d'amis, l'Audience conditionnant ses décisions à l'observation des décisions des premiers membres. Cela implique qu'en faisant de fausses déclarations, le premier ami nuit à l'image du second et réduit ses opportunités économiques. Cependant, nous ne trouvons aucune preuve que les retombées potentielles du comportement du premier ami sur l'image sociale du deuxième ami dissuadent les fausses déclarations. Les individus n'intériorisent pas l'effet de leur comportement sur l'image de leurs amis. Nous explorons plusieurs explications possibles, mais la plus convaincante est que les membres des paires sous-estiment l'impact de leur comportement sur la décision de l'Audience.

Le chapitre 2 contribue à une littérature très limitée en économie qui étudie comment les groupes construisent leurs images sociales (Tirole, 1996; Huck and Lünser, 2010). En accord avec cette littérature, nous constatons que l'image du groupe a des caractéristiques de bien public, parce que les tierces personnes utilisent l'information sur le comportement des membres qu'ils peuvent observer pour faire des inférences à propos des autres membres. Mais la nouveauté est que les membres du groupe n'intériorisent pas l'effet de leur comportement sur l'image du groupe non pas parce qu'ils ne se soucient pas des membres du groupe, mais plutôt parce qu'ils n'anticipent pas que leur comportement est utilisé pour tirer des conclusions sur le groupe dans son ensemble.

Ceci met en évidence un hiatus intéressant entre les perceptions des tierces personnes, qui perçoivent les membres comme homogènes, et les perceptions des membres du groupe, qui ne se considèrent pas comme représentatifs de leur groupe.

Préférences sociales :

Les "modèles formels de préférences sociales... supposent que les gens s'intéressent à leurs propres intérêts, mais qu'ils sont aussi préoccupés par les gains des autres" (Charness and Rabin, 2002). En tant que telles, les préférences sociales sont un moyen très général de prendre en compte les motivations sociales dans les décisions économiques. La littérature sur les préférences sociales a pris de l'ampleur grâce aux résultats empiriques issus de jeux de négociation simples difficilement rationalisables sous l'hypothèse d'égoïsme (Güth et al., 1982). Dans ces jeux, deux joueurs négocient le partage d'une dotation. Le premier joueur propose un partage, et le second décide d'accepter ou non ce partage. Si le deuxième joueur rejette l'offre, les deux joueurs ne gagnent rien. L'équilibre de Nash standard d'un tel jeu est pour le premier joueur d'envoyer un montant très faible, et pour le deuxième joueur d'accepter parce que même ce petit montant est préférable à ne rien gagner. Empiriquement, les propositions des premiers joueurs sont souvent substantielles, tandis que les seconds rejettent généralement les offres trop faibles. L'interprétation donnée est que les seconds joueurs refusent les offres basses parce qu'ils pensent que les partages ainsi implémentés seraient injustes. Les premiers joueurs anticipent ces préoccupations en matière d'équité et proposent des montants non négligeables. Les premiers joueurs peuvent également prendre en compte le profit du second joueur indépendamment du risque de rejet de l'offre.

Un certain nombre de motivations sociales sont contenues dans le concept de préférences sociales. L'altruisme, qui a été identifié dans des décisions simples telles que le jeu du dictateur (Forsythe et al., 1994; Hoffman et al., 1994) ou les contributions dans les jeux de bien public (Andreoni, 1995), est la propension à partager les ressources

de manières désintéressée. Le souci d'équité est la raison pour laquelle les individus sacrifient des gains monétaires en rejetant les petites offres afin d'éviter des résultats injustes (Güth et al., 1982). L'aversion pour l'inégalité reflète le fait que certains individus n'aiment pas faire face à des inégalités trop marquées, y compris quand ces inégalités sont à leur avantage (Fehr and Schmidt, 1999; Bolton and Ockenfels, 2000). La réciprocité désigne le fait que la préoccupation d'un décideur vis-à-vis de la situation d'une tierce personne dépend de comment cette tierce personne s'est comportée vis-à-vis du décideur auparavant (Charness and Rabin, 2002). Les préférences sociales ne sont pas nécessairement nobles, car l'intérêt porté à la situation d'autrui peut aussi parfois susciter de la jalousie ou de la malveillance.

La prise en compte des préférences sociales dans la fonction d'utilité a permis d'expliquer un certain nombre de questions économiques importantes. Les considérations d'équité génèrent une certaine rigidité des prix (Fehr et al., 1993), et peuvent donc être responsables de la rigidité des salaires et causer du chômage. L'altruisme favorise la fourniture de biens publics (Andreoni, 1995; Goeree et al., 2002). La réciprocité facilite la coopération (Fehr et al., 2002).

Le chapitre 3 contribue à la littérature sur les préférences sociales, et plus particulièrement sur les déterminants de l'altruisme. Dans notre expérience de laboratoire, les participants doivent d'abord faire une tâche d'effort réel pour un salaire fixe, puis jouer à un jeu du dictateur avec une dotation de €5. Dans un design 3×2 , nous faisons varier le niveau du salaire pour la tâche d'effort réel entre trois niveaux (faible = 5 euros, moyen = 10 euros et élevé = 15 euros) et le moment où le salaire est révélé. Dans certaines conditions, les participants connaissent le salaire qu'ils gagnent pour la tâche d'effort réel avant de l'effectuer, et ne connaissent pas les autres niveaux potentiels du salaire. De plus, ils reçoivent une enveloppe contenant leurs salaires avant la tâche. Dans d'autres conditions, les participants sont informés de la distribution des salaires potentiels avant la tâche d'effort réel, mais ils ne sont informés du vrai montant de leur salaire qu'après

la tâche. Les trois niveaux de salaire sont également probables. Dans ces conditions, nous donnons aux participants une enveloppe contenant le salaire moyen avant la tâche d'effort réel, de laquelle nous retirons de l'argent si le vrai salaire s'avère être le salaire faible ou nous rajoutons de l'argent si le vrai salaire s'avère être le salaire élevé. Notre hypothèse est que les participants dans ces dernières conditions évaluent le salaire qu'ils reçoivent en effet en comparaison des autres niveaux qu'ils auraient pu recevoir. Dans les conditions décrites plus haut en revanche, ils ne peuvent pas comparer leur salaire effectif aux autres niveaux potentiels. Nous recueillons également des mesures autodéclarées et physiologiques sur les émotions ressenties par les participants au cours de l'expérience, afin d'étudier le rôle des émotions comme la déception ou l'exaltation sur les transferts ultérieurs.

Pour formaliser comment l'évaluation du salaire relativement aux autres niveaux potentiels peut affecter la générosité dans le jeu du dictateur, nous développons un modèle de partage qui présente "de la dépendance par rapport aux références" (*reference dependence*) (Kahneman and Tversky, 1979; Tversky and Kahneman, 1991; Köszegi and Rabin, 2006). Dans le modèle, les participants forment une référence sur ce qu'ils gagneront dans l'expérience lorsqu'ils reçoivent les instructions décrivant l'expérience. Cette référence dépend du salaire de référence et du transfert hypothétique qu'ils choisiraient si le salaire vrai s'avérait égal au salaire de référence. Le salaire de référence est supposé se situer entre le salaire le plus élevé et le salaire le moins élevé, de sorte que le salaire le plus élevé est encodé comme un gain et le salaire le moins élevé comme une perte. Cette hypothèse est conforme à la formation du point de référence fondée sur le statu-quo et sur les attentes. De plus, les participants ne peuvent pas formuler de façon plausible un point de référence inférieur au salaire faible ou supérieur au salaire élevé. Cela garantit qu'au moins certains participants encodent le salaire faible (resp. élevé) comme une perte (resp. gain) alors que d'autres ne l'encodent ni comme une perte ni comme un gain. Le modèle prévoit que, pour les participants qui reçoivent le salaire faible, ceux qui sont informés des autres niveaux de salaire potentiels transfèrent moins que les participants

qui ne sont pas informés des autres niveaux de salaire potentiels. L'interprétation est que la réduction de leur transfert dans le jeu du dictateur leur permet de minimiser la différence négative entre leur référence et leurs gains réels dans l'expérience. Le modèle a une prédiction opposée pour le cas où les participants reçoivent le salaire élevé. La première prédiction n'est vérifiée que pour les participants qui sont averses aux pertes. Lorsque les participants sont averses aux pertes, les informer qu'ils auraient pu obtenir un salaire plus élevé réduit leur préoccupation pour autrui. D'autre part, la seconde prédiction est vérifiée indépendamment de l'aversion pour les pertes des participants. En ce qui concerne les émotions : les participants éprouvent des émotions de la valence attendue lorsque le salaire vrai est révélé, mais les émotions ont très peu de pouvoir explicatif sur les transferts. Cela suggère que nos résultats saisissent un effet distinct de l'effet des émotions sur l'altruisme.

Nous manipulons le contexte en faisant varier l'information concernant les salaires potentiels pour une tâche d'effort réel, et nous montrons que cela a un impact sur la générosité dans un jeu du dictateur ultérieur. Ce chapitre appartient donc à la littérature sur les déterminants contextuels de l'altruisme. Les résultats vont dans le sens de diverses études montrant que de subtiles manipulations du contexte dans lequel la décision est prise influencent substantiellement l'altruisme en changeant la norme perçue (List, 2007; Bardsley, 2008), en fournissant des excuses pour ne pas donner (Dana et al., 2006) ou en réaffectant les droits de propriété (Cherry et al., 2002) par exemple. Une conclusion prématurée pourrait être que les préférences sociales sont par nature très erratiques, ce qui mettrait en doute leur pouvoir prédictif. Une autre conclusion est que les préférences sociales sont plutôt stables chez les individus, mais que la compréhension des comportements motivés par l'altruisme, le souci d'équité ou l'aversion pour l'inégalité exige un examen approfondi des circonstances des décisions. Cet examen peut à son tour fournir des informations intéressantes sur les motivations sociales et psychologiques des individus (Falk and Heckman, 2009).

Bibliography

Abeler, J., D. Nosenzo, and C. Raymond (2018). Preferences for truth-telling. *Econometrica*.

Akerlof, G. and J. Yellen (1988). Fairness and unemployment. *American Economic Review 78*(2), 44–49.

Akerlof, G. A. (1982). Labor contracts as partial gift exchange*. *The Quarterly Journal of Economics 97*(4), 543–569.

Akerlof, G. A. and J. L. Yellen (1990). The fair wage-effort hypothesis and unemployment. *The Quarterly Journal of Economics 105*(2), 255–283.

Andreoni, J. (1995). Cooperation in public-goods experiments: kindness or confusion? *The American Economic Review*, 891–904.

Andreoni, J. and B. D. Bernheim (2009). Social image and the 50-50 norm: A theoretical and experimental analysis of audience effects. *Econometrica 77*(5), 1607–1636.

Andreoni, J. and R. Petrie (2004). Public goods experiments without confidentiality: a glimpse into fund-raising. *Journal of Public Economics 88*(7), 1605 – 1623.

Andreoni, J., J. M. Rao, and H. Trachtman (2017). Avoiding the ask: A field experiment on altruism, empathy, and charitable giving. *Journal of Political Economy 125*(3), 625–653.

Ariely, D., A. Bracha, and S. Meier (2009, March). Doing good or doing well? image motivation and monetary incentives in behaving prosocially. *American Economic Review 99*(1), 544–55.

Arrow, K. (1973). The theory of discrimination. *Discrimination in labor markets 3*(10), 3–33.

Ashraf, N., C. F. Camerer, and G. Loewenstein (2005, September). Adam smith, behavioral economist. *Journal of Economic Perspectives 19*(3), 131–145.

Bardsley, N. (2008, Jun). Dictator game giving: altruism or artefact? *Experimental Economics 11*(2), 122–133.

Becker, G. S. (1974). A theory of social interactions. *Journal of Political Economy 82*(6), 1063–1093.

Bénabou, R. and J. Tirole (2006, December). Incentives and prosocial behavior. *American Economic Review 96*(5), 1652–1678.

Benjamin, D. J., J. J. Choi, and A. J. Strickland (2010). Social identity and preferences. *American Economic Review 100*(4), 1913–28.

Berg, J., J. Dickhaut, and K. McCabe (1995). Trust, reciprocity, and social history. *Games and Economic Behavior 10*(1), 122 – 142.

Berman, E. (2000). Sect, subsidy, and sacrifice: An economist's view of ultra-orthodox jews. *The Quarterly Journal of Economics 115*(3), 905–953.

Bernheim, B. D. (1994). A theory of conformity. *Journal of Political Economy 102*(5), 841–877.

Bolton, G. E. and A. Ockenfels (2000). Erc: A theory of equity, reciprocity, and competition. *American economic review 90*(1), 166–193.

Brandts, J. and N. Figueras (2003). An exploration of reputation formation in experimental games. *Journal of Economic Behavior & Organization 50*(1), 89 – 115.

Bursztyn, L. and R. Jensen (2015). How does peer pressure affect educational investments? *The Quarterly Journal of Economics 130*(3), 1329–1367.

Camerer, C. and K. Weigelt (1988). Experimental tests of a sequential equilibrium reputation model. *Econometrica 56*(1), 1–36.

Carvalho, J.-P. (2013). Veiling. *The Quarterly Journal of Economics 128*(1), 337–370.

Cettolin, E. and S. Suetens (2018). Return on trust is lower for immigrants. *The Economic Journal*.

Charness, G., D. Masclet, and M. C. Villeval (2014). The dark side of competition for status. *Management Science 60*(1), 38–55.

Charness, G. and M. Rabin (2002). Understanding social preferences with simple tests. *The Quarterly Journal of Economics 117*(3), 817–869.

Charness, G., L. Rigotti, and A. Rustichini (2007, September). Individual behavior and group membership. *American Economic Review 97*(4), 1340–1352.

Chen, R. and Y. Chen (2011, October). The potential of social identity for equilibrium selection. *American Economic Review 101*(6), 2562–89.

Chen, Y. and S. X. Li (2009, March). Group identity and social preferences. *American Economic Review 99*(1), 431–57.

Cherry, T. L., P. Frykblom, and J. F. Shogren (2002). Hardnose the dictator. *The American Economic Review 92*(4), 1218–1221.

Cohn, A., M. A. Maréchal, and T. Noll (2015). Bad boys: How criminal identity salience affects rule violation. *The Review of Economic Studies 82*(4), 1289–1308.

Dana, J., R. a. Weber, and J. X. Kuang (2006, sep). Exploiting moral wiggle room: experiments demonstrating an illusory preference for fairness. *Economic Theory 33*(1), 67–80.

Duesenberry, J. (1949). *Income: Saving, and the Theory of Consumer Behavior*. Harvard economic studies. HarvardUniversity Press.

Falk, A. and J. J. Heckman (2009). Lab experiments are a major source of knowledge in the social sciences. *science 326*(5952), 535–538.

Fehr, E., U. Fischbacher, and S. Gächter (2002). Strong reciprocity, human cooperation, and the enforcement of social norms. *Human nature 13*(1), 1–25.

Fehr, E., G. Kirchsteiger, and A. Riedl (1993). Does fairness prevent market clearing? an experimental investigation. *The Quarterly Journal of Economics 108*(2), 437–459.

Fehr, E. and K. M. Schmidt (1999). A theory of fairness, competition, and cooperation. *The quarterly journal of economics 114*(3), 817–868.

Fischbacher, U. and F. Föllmi-Heusi (2013). Lies in disguiseâ€"an experimental study on cheating. *Journal of the European Economic Association 11*(3), 525–547.

Forsythe, R., J. L. Horowitz, N. Savin, and M. Sefton (1994). Fairness in simple bargaining experiments. *Games and Economic Behavior 6*(3), 347 – 369.

Friedman, M. (1953). *Essays in Positive Economics*. A Phoenix book. Business economics. University of Chicago Press.

Gneezy, U. (2005). Deception: The role of consequences. *American Economic Review 95*(1), 384–394.

Goeree, J. K., C. A. Holt, and S. K. Laury (2002). Private costs and public benefits: unraveling the effects of altruism and noisy behavior. *Journal of Public Economics 83*(2), 255 – 276.

Güth, W., R. Schmittberger, and B. Schwarze (1982). An experimental analysis of ultimatum bargaining. *Journal of Economic Behavior & Organization 3*(4), 367 – 388.

Harbaugh, W. T. (1998). The prestige motive for making charitable transfers. *The American Economic Review 88*(2), 277–282.

Hargreaves Heap, S. P. and D. J. Zizzo (2009, March). The value of groups. *American Economic Review 99*(1), 295–323.

Hedegaard, M. S. and J.-R. Tyran (2018, January). The price of prejudice. *American Economic Journal: Applied Economics 10*(1), 40–63.

Hoffman, E., K. McCabe, K. Shachat, and V. Smith (1994). Preferences, property rights, and anonymity in bargaining games. *Games and Economic behavior 7*(3), 346–380.

Huck, S. and G. K. Lünser (2010). Group reputations: An experimental foray. *Journal of Economic Behavior & Organization 73*(2), 153–157.

Huck, S., G. K. Lünser, and J.-R. Tyran (2012). Competition fosters trust. *Games and Economic Behavior 76*(1), 195–209.

Kahneman, D., J. L. Knetsch, and R. Thaler (1986). Fairness as a constraint on profit seeking: Entitlements in the market. *The American economic review*, 728–741.

Kahneman, D. and A. Tversky (1979). Prospect theory: An analysis of decision under risk. *Econometrica 47*(2), 263–291.

Kajackaite, A., U. Gneezy, and J. Sobel (2018). Lying aversion and the size of the lie. *The American Economic Review*.

Köszegi, B. and M. Rabin (2006). A model of reference-dependent preferences. *The Quarterly Journal of Economics 121*(4), 1133–1165.

Lane, T. (2016). Discrimination in the laboratory: A meta-analysis of economics experiments. *European Economic Review 90*, 375 – 402. Social identity and discrimination.

List, J. A. (2007). On the Interpretation of Giving in Dictator Games. *Journal of Political Economy 115*(3), 482–493.

Montmarquette, C., J.-L. Rullière, M.-C. Villeval, and R. Zeiliger (2004). Redesigning teams and incentives in a merger: An experiment with managers and students. *Management Science 50*(10), 1379–1389.

Phelps, E. S. (1972). The statistical theory of racism and sexism. *The american economic review*, 659–661.

Riley, J. G. (2001). Silver signals: Twenty-five years of screening and signaling. *Journal of Economic literature 39*(2), 432–478.

Shariff, A. F. and A. Norenzayan (2007). God is watching you: Priming god concepts increases prosocial behavior in an anonymous economic game. *Psychological science 18*(9), 803–809.

Smith, A. (1759). *The Theory of Moral Sentiments*. McMaster University Archive for the History of Economic Thought.

Smith, V. L. (2012). Adam smith on humanomic behavior. *The Journal of Behavioral Finance & Economics 2*(1), 1–20.

Smith, V. L. (2013). Adam smith: From propriety and sentiments to property and wealth. *Forum for Social Economics 42*(4), 283–297.

Spence, M. (1973). Job market signaling. *The Quarterly Journal of Economics 87*(3), 355–374.

Tajfel, H. and J. C. Turner (1979). An integrative theory of intergroup conflict. *The social psychology of intergroup relations 33*(47), 74.

Tirole, J. (1996). A theory of collective reputations (with applications to the persistence of corruption and to firm quality). *The Review of Economic Studies 63*(1), 1–22.

Tversky, A. and D. Kahneman (1991). Loss aversion in riskless choice: A reference-dependent model. *The Quarterly Journal of Economics 106*(4), 1039–1061.

Wilson, R. (1985). Reputations in games and markets. *Game-theoretic models of bargaining*, 27–62.

Contents

List of Tables

List of Figures

Introduction

This thesis studies three social determinants of economic decisions: Social Identity, Social Image and Social Preferences. Individuals make decisions in social contexts and economics, a social science of choices, has become more and more interested in understanding how social factors influence economic decisions.

In the seminal work of Adam Smith (1759) altruism, fairness or social esteem were already identified among the determinants of economic decisions (Ashraf et al., 2005): in some circumstances, individual behavior are not ruled by passions, which naturally lead to selfishness, but by the "impartial spectator" who seeks fitness to the social rules. Vernon Smith (2012, 2013) refers to the deep insights contained in the seminal work of Adam Smith by the word "Humanomics". The bottom line is that a lot of economic behavior remain cryptic if we ignore that individuals seek praiseworthiness and flee blameworthiness, and if they are not analyzed in light of the emerging social rules that define which behavior are blameworthy and which behavior are praiseworthy.

However, economics has somewhat relegated these motivations to the background. In his early description of the Economic Man, John Stuart Mill outlined the motivations behind economic decisions as material, selfish, and asocial (Mill, 1874). The simple figure of the Economic Man has dominated the conceptual universe of economists for a long time, because it constituted a simple and powerful framework which provided important insights on a variety of economic questions. Social motivations may also

be less amenable to reductionism, an important pillar of the methodology of positive economics (Friedman, 1953), which might explain why this type of social motivations remained considered as of second order in the understanding of economic choices. In other words, the Economic Man has survived despite of its weaknesses because the analyze of his choices is sufficient to explain a large number of economic behavior in a clear and compelling way.

Besides his weaknesses and his apparent asociality, the Economic Man presents great strengths for explaining social behavior in a parsimonious fashion. First, the starkness of the assumption about its motivations allows a neat study of social interactions. Game theory, the study of strategic interactions, formalizes why individual decisions cannot be understood with at least some account for the social context. A self-interested player who interacts with other self-interested players have to take into account the other players' preferences, because his optimal strategy depends on the others' strategies which in turn depends on their preferences.

In addition, the economic calculus supposed to drive the Economic Man's decisions can be applied to (supposedly) non-economic decisions or *behavior*. As such, Gary Becker's economics of family is an important milestone in the history of economic thought. Becker (1974) endowed the Economic Man with some form of "social" motivation: in his theory, the Economic Man takes into account the welfare of his offspring in his economic calculus. Altruism is then recognized as a legitimate motivation of the Economic Man, as for example bequests are explained by the concern of the family head for the welfare of the heirs. However, Becker's account for altruism can be considered as minimal because it is assumed to influence behavior in a very restricted set of decisions: individuals may be altruistic in the family, but selfishness is still a premise for explaining market interactions.

The endeavour of integrating social elements in the motivation of the Economic Man has provided new perspectives on old economic problems. For instance, Duesenberry (1949) proposed that consumption fulfills a social role: conspicuous consumption provides positional status. This can explain why aggregate savings do not increase as expected when the aggregate income increases. Akerlof (1982) and Akerlof and Yellen (1988, 1990) imported insights from social psychology to explain wage rigidity and unemployment: they show that if workers are concerned by the fairness of their wage, firms cannot adjust wages downward in case of economic slowdown and this leads to unbalances on the labor market. The concern for fairness, a social motivation, became a compelling antecedent of unemployment. This demonstrate that social motivations can have important implications even in the market.

More recently, an accumulation of data from lab and field experiments brought to the fore a number of empirical findings that are hardly compatible with the hypothesis of the Economic Man, and more crucially show the limits of its predictive power. For instance: Why don't monopolistic firms utilize their bargaining power to fix high prices (Kahneman et al., 1986; Fehr et al., 1993)? Why don't people fully exploit opportunities to lie or cheat and leave money on the table (Gneezy, 2005; Fischbacher and Föllmi-Heusi, 2013)? Why do people give money to charities when asked but often use costly strategies to avoid the ask (Andreoni et al., 2017)? Why do people sometimes act less prosocially when they have extrinsic motivations to do so than when they do not (Gneezy and Rustichini, 2000; Bénabou and Tirole, 2006; Ariely et al., 2009)? Why do people invest costly efforts to achieve statuses that have no economic payoff (Charness et al., 2014)? Friedman (1953) argued that the quality of a theoretical hypothesis is independent of its realism but is determined by its capacity to predict behavior. In very simple decisions, for instance whether to cheat for money, the hypothesis of the Economic Man are unable to predict behavior. A minutious examination of the relevance of additional or alternative hypotheses about what drives economic behavior is necessary to improve the ability of economic theory to understand economic phenomena. Among these hypotheses, we

will focus on those related to social motivations.

In this thesis, we study social identity, social image and social preferences, three social motivations of interest to economics because they help understand the empirical regularities presented in the previous paragraph. For instance, the willingness to preserve a good social image partially explains why people do not fully exploit cheating opportunities: some people remain honest to keep the esteem of others. Accounting for social identity illuminates why people are willing to sacrifice resources to achieve a high status which provides no economic payoff: it allows them to define *who they are* in a positive way. Social preferences may explain why firms do not exploit their bargaining power: customers who dislike being treated unfairly would boycott them. The following sections review the literature on social identity, social image and social preferences in economics in more details and present our contributions to each of these literatures.

Social identity

Identity is a person's sense of self (Akerlof and Kranton, 2000), or in other terms the answer one gives to the question "'Who Am I?". Individuals seek to define and signal their identity with their behavior. As a consequence, to fully understand economic decisions, it is important to account for the meaning individuals try to convey with their behavior and how it contributes to build identity. For instance, Bénabou and Tirole (2006) explain that acting prosocially is a way to signal virtue and contributes to define one's identity. They present an interesting implication of the former: extrinsic motivations might reduce the willingness to act prosocially, because they blur the signal about one's intrinsic motivations. Taking motivations related to identity into account improves the understanding of the power and limitations of incentives.

An important part of any person's identity is derived from group membership, which constitutes Social Identity (Tajfel and Turner, 1979). Identification to social groups im-

pacts behavior substantially. A striking example is the ease with which researchers can generate animus-based discrimination in the laboratory: many economic experiments arbitrarily separate the participants into salient groups, and find ingroup bias, i.e. a tendency of participants to favor those with whom their share a common group identity (e.g. Chen and Li, 2009; Hargreaves Heap and Zizzo, 2009). Lab induced group identities also impact strategic behavior in coordination and cooperation games (see e.g. Charness et al., 2007; Chen and Chen, 2011). The results with lab induced group identities are echoed by studies using naturally occurring groups: for instance, ingroups bias is found between groups delineated by the boundaries of firms (Montmarquette et al., 2004), religion (e.g. Hedegaard and Tyran, 2018) or ethnicity (e.g Cettolin and Suetens, 2018). It is interesting to note that natural groups do not generally induce more ingroup bias than lab induced groups (Lane, 2016).

Social identity have economic implications beyond the mere ingroup bias. By priming the membership in groups, researchers find that the members tend to conform to the stereotypes of the group. For instance, when their group membership is made salient, Asian-American become more patient (Benjamin et al., 2010), inmates become more dishonest (Cohn et al., 2015) and religious people become more generous (Shariff and Norenzayan, 2007).

Chapter 1 shows that the concept of social identity provides interesting insights on an important economic issue. Our original research question is to study how social mobility, defined as the movement from a social level to a higher one, impacts trust and trustworthiness in the presence of group identity. A concrete example of upward mobility is promotion within firms. Promotions have motivational effects, and the lack thereof might reduce the incentives of workers to provide efforts. However, promotions may interact with social identity: our premise is that promotion increases the social distance between the promoted and the left-behinds, while it might decrease the social distance between the promoted and out-groups who already have the high status. We

conjecture that upward mobility might impact trust through social distance. The effect of upward mobility on trust is of practical relevance, because trust between co-workers is a substitute for costly monitoring and can mitigate the problem of incomplete contracts, and as such constitutes an asset for the success of firms.

In order to measure the effects of upward mobility on trust, we experimentally studied a standard trust game (Berg et al., 1995) in which players were identified by a natural group identity and a status. The group identity was conferred by the school of the subjects: either the local business school or the local engineering school. Both schools are selective and have specific cultural identities. As a consequence, students identify strongly with their schools. The natural identity of subjects was reinforced in the first part of the experiment by means of a quiz about their respective schools.

In contrast, status was induced in the laboratory. Subjects were asked to answer to a math quiz. Relative performance in this task determined the role of subjects in the last part of the experiment. We called "Experts" the subjects who gave the higher number of correct answers, and "Agents" the remaining subjects. This introduced a hierarchy between roles. A critical feature of the experimental design is that school affiliation predicts the performance in the math quiz. As a consequence, schools predict the status of subjects, but imperfectly. Subjects from the engineering school were much more likely to achieve the high status, and we represent upward mobility by assigning the status of expert also to the best performers from the business school. This captures an important feature of upward mobility, as these subjects achieved a high status that was ex-ante unlikely given their school. Importantly, our design disconnected status from payoffs: experts and agents have different roles but receive the same payoffs. This isolates the pure effect of status and identity on trust and trustworthiness.

Then, subjects played a trust game under the strategy method. They made decisions both as trustor and as trustee, being informed of the school and the status of expert or

agent of their counterpart in all possible combinations. In order to measure the impact of upward mobility on trust and trustworthiness, we compared the transfers of experts from the business school to the transfers of agents from the business school. We also compared the transfers directed at experts from the business school to those directed at agents from the business school in order to measure the impact of upward mobility on the propensity of others to trust and to be trustworthy to upwardly mobile subjects.

The results show that upward mobility does impact trust, in a specific way: upwardly mobile subjects tend to trust less, irrespective of the identity of the trustee. On the other hand, upwardly mobile individuals do not attract discriminatory behavior: they receive the same transfers as other subjects from the business school. These results, taken together, suggest that the effect of upward mobility on social distance is subtle. Upwardly mobile individuals seem to distantiate from both ingroups and outgroups, which can be explained by the specific experience of having the high status while being from the business school: experts from the business school might perceive that their status is more rewarding for them than for the subjects from the engineering school because they were unlikely to reach it ex-ante. Hence, they might consider themselves as "singled-out": their status provides them a specific identity that is not a simple combination of a status and school.

Chapter 1 also provides a more fundamental contribution to the literature on social identity. The results indicate that a status induced in the lab, which has no monetary payoff and last only for the length of the experiment, can undo a natural identity. This confirms that status is a strong source of identity.

Social image:

The second social determinant that is studied in this thesis is social image, i.e. the perception of intrinsic characteristics of an individual by an audience. The economic

approach to social image is linked to the development of information economics. In every economic interaction in which information about the stake-holders is imperfect, perceptions about the stake-holders is substituted to the unavailable accurate information. Beliefs-based discrimination (Phelps, 1972; Arrow, 1973), in which the social image of individuals is fed by stereotypes about the group to which they belong, illustrates clearly why social image has important economic implications: those who are stereotyped-against incur large economic costs.

Social image enters the utility function in two distinct ways. First, it impacts utility indirectly, through the impact of social image on interactions with others. Individuals take costly actions to signal their inner types (see Riley, 2001, for a review) and to preserve their reputations (see Wilson, 1985, for a review). Students get diplomas (in part) to be perceived as competent by employers (Spence, 1973) or conversely neglect their education to demonstrate their conformity to a culture of low educational achievement (Bursztyn and Jensen, 2015). Religious people adopt economically costly customs to signal the strength of their commitment to the cult, and therefore benefit from interactions with its members (Berman, 2000; Carvalho, 2013). When information about past behavior is available, untrustworthy individuals mimic trustworthy ones in order to preserve their chances of being trusted in later interactions (Camerer and Weigelt, 1988; Brandts and Figueras, 2003; Huck et al., 2012).

Social image also enters the utility function directly as many people care about how they are perceived by others *per se*. In theoretical models, the taste for social esteem have been used to explain pro-social behavior (Harbaugh, 1998; Bénabou and Tirole, 2006; Andreoni and Bernheim, 2009), norm obedience (Bernheim, 1994) or honesty (Abeler et al., 2018; Kajackaite et al., 2018), among others. The empirical literature, and in particular experiments varying the level of observability of non-strategic decisions, clearly supports the hypothesis that individuals attach an important weight to the perceptions of others, even when these perceptions cannot have material consequences. Hoffman

et al. (1994) show that using a procedure in which neither the experimentalist nor the recipients could identify the dictator reduces transfer drastically compared to a procedure in which only the experimentalists could identify the dictator. Andreoni and Petrie (2004) vary whether pictures are used to identify participants in a public goods experiment, and find that provision is greater when pictures are used. In a real-effort task in which performances determined payment to a charity, Ariely et al. (2009) show that participants have higher performances when they have to publicly reveal their performance to other participants than when the performance remains private. Fischbacher and Föllmi-Heusi (2013) ask participants to report the outcome of a die-roll that they privately observe and this report conditions their payoffs. The distribution of reports shows that there is partial lying, which indicates that participants adopt sophisticated strategies to avoid being perceived as dishonest.

Chapter 2 provides an original contribution to the economics of social image. Social image is generally considered as a "private good": individuals act more pro-socially, or more honestly to be perceived positively *as individuals*. The novelty of this chapter is to experimentally study a set-up in which social image has some public good properties: individual behavior might spill over onto the in-groups' social image. The game that we experimentally study is a three-players game. Two players are members of a pair, and the third one is called "the Audience". First, both members of the pair report the outcome of a random variable independently and without communication. They sometimes have monetary incentives to misreport. After both members of the pair have made their report, the Audience observes the report of one predetermined member of the pair and the actual value of the random variable, so that he can spot misreporting. Then, the Audience must decide whether he wants to play with the second member of the pair. The payoff of the audience is such that he wants to play only if he expects that the probability of a truthful report by the second member is high enough. Both members of the pair prefer the Audience to play, because when the Audience plays, they get an extra payoff. Between conditions, we vary whether the behavior of the first member

of the pair is informative of the behavior of the second member, which is important for the Audience who has to form a belief about the second member of the pair. In the baseline condition, the members of the pair do not know each other. In the treatment, the members of the pair are friends, which makes the behavior of the first member potentially informative about the behavior of the second. Indeed, the decisions of two friends are correlated since friends tend to cluster by preference because of homophily and mutual influence.

The data show that social image is indeed a public good in pairs of friends, as audience condition their decisions on the observation of the first members' decisions. This means that, by misreporting, the first friends harm the image of the second one and reduces his economic opportunities. However, we find no evidence that the potential spillovers from the behavior of the first friend to the social image of the second friend deter misreporting. Individuals do not internalize the effect of their behavior on the image of their friends. We explore several potential explanations, but the most compelling one is that members of pairs underestimate the impact of their behavior on the decision of the Audience.

Chapter 2 contributes to a very scarce literature in economics studying how groups build their social images (Tirole, 1996; Huck and Lünser, 2010). In line with this literature, we find that the image of the group has public good features, because outsiders uses the information about the behavior of the members they can observe to form inferences about the other members. But the novelty is that group members do not internalize the effect of their behavior on the image of the group not because they do not care about the ingroups, but rather because they do not anticipate that their behavior is used to form inferences about the group as a whole. This points to an interesting gap between the perceptions of outsiders, who perceive the members has homogeneous, and the perceptions of the ingroups, who consider that they are not representative of their group. Why this gap exists and how it can be reduced are interesting questions for further

research.

Social preferences:

The "formal models of social preferences...assume people are self-interested, but are also concerned about the payoffs of others" (Charness and Rabin, 2002). As such, social preferences are a most general way of accounting for social motivations in economic decisions. The literature on social preferences have gain momentum thanks to empirical findings in simple bargaining games that were hardly rationalizable under the hypothesis of selfishness (Güth et al., 1982). In these games, two players bargain over an endowment. The first mover proposes a sharing, and the second mover decides to accept this offer or not. If the second mover rejects the offer, both players earn zero. The standard Nash equilibrium of such a game is for the first player to send a very low amount, and for the second player to accept because even this small amount is better than zero. Empirically, the proposition of first movers are often substantial, while the second movers generally reject low offers. The interpretation of the deviations from the Nash Equilibrium is that second movers refuse low offers because they think that the resulting outcome would be unfair. First movers anticipate these fairness concerns and propose non-trivial amounts. First movers can also care for the payoff of the second mover independently of the risk rejection.

A number of social motivations are nested in the concept of social preferences. Altruism, which was identified in simple sharing decisions such as dictator games with tight control on observability (Forsythe et al., 1994; Hoffman et al., 1994) or in public good contributions (Andreoni, 1995), is the selfless inclination to share resources. Fairness concerns is the reason why individuals sacrifice monetary gains by rejecting small offers in order to avoid unfair outcomes (Güth et al., 1982). Inequity aversion captures the fact that some individuals dislike facing inequality, even advantageous ones (Fehr

and Schmidt, 1999; Bolton and Ockenfels, 2000). Reciprocity points to the fact that the concern for the payoff of a given individual depends on whether this individual acted kindly or not toward the decision maker (Charness and Rabin, 2002). Social preferences are not necessarily noble, as the concern for the payoff of others also may sometimes generates envy or spite.

Accounting for social preferences in the utility function has been able to explain a number of important economic questions. Fairness considerations generate some price rigidity (Fehr et al., 1993), and therefore may be responsible of fairness-based wage rigidity that causes unemployment. Altruism promotes the provision of public goods (Andreoni, 1995; Goeree et al., 2002). Reciprocity facilitates cooperation (Fehr et al., 2002).

Chapter 3 contributes to the literature on social preferences, and more specifically on what determines altruism. In our laboratory experiment, participants first have to perform a real effort task for a fixed wage, and then play a dictator game with a €5 endowment. In a 3×2 design, we vary the level of the wage for the real effort task between three levels (low =€5, medium=€10 and high =€15) and the timing of revelation of the wage. In some conditions, participants know the actual wage for the real effort task before performing the real effort task, and are not aware of the other potential levels. Moreover, they are given an envelope containing their wage before the real effort task. In some other conditions, participants are informed of the distribution of the potential wages before the real effort task, but they are informed of their actual wage only after the real effort task. The three levels of wage are equally likely. In these conditions, they are given an envelope containing the medium wage before the real effort task, from which we take out money if the actual wage is low or add money if the actual wage is high. Our premise is that participants in the latter conditions evaluate their actual wage relative to the other levels they could have received. In the former condition, they cannot compare their actual wage to the other potential levels. We also collect self-reported and physiological measures of emotions, to investigate the role of

emotions such as disappointment or elation on subsequent transfers.

To formalize how evaluating the wage relative to other potential levels can affect generosity in the subsequent dictator game, we develop a model of sharing that presents reference-dependence (Kahneman and Tversky, 1979; Tversky and Kahneman, 1991; Köszegi and Rabin, 2006). In the model, participants form a reference about how much they will earn in the experiment when they receive the instructions describing the experiment. This reference depends on the reference wage and the hypothetical transfer they would choose if the actual wage is equal to the reference wage. The reference wage is assumed to be in-between the high and the low wage, so that the high wage is encoded as a gain while the low wage is encoded as a loss. This assumption is consistent with both the status-quo based and the expectation-based formation of the reference point, respectively because the envelops contain the medium wage and because the expected wage is equal to the medium wage. Moreover, participants cannot plausibly formulate a reference point below the low wage or above the high wage. This insures that at least some participants encode the low (resp. high) wage as a loss (resp. gain) while the others encode it neither as a loss nor as a gain. The model predicts that, for participants who get the low wage, those who are informed of the other potential levels of wage transfer less than participants who are not informed of the other potential levels. The interpretation is that reducing their transfers in the dictator game allows them to minimize the negative difference between their reference and their actual earnings in the experiment. The model has an opposite prediction for the case in which participants receive the high wage. The former prediction is supported by the data for loss averse participants only. When participants are loss averse, informing them that they could have gotten a higher wage reduces their concern for the recipient. On the other hand, the latter prediction is supported irrespective of the loss aversion of participants. Regarding emotions: participants do experience emotions of the expected valence when the actual wage is revealed, however emotions have very little explanatory power on transfers. This suggests that our results capture an effect that is distinct from the effect of emotions

on altruism.

Because we manipulate the context by varying the salience of counter-factual wages for a real effort task, and show that this impacts generosity in a subsequent dictator game, this chapter belongs to the literature on the contextual determinants of altruism. The results are in line with various studies showing that subtle manipulations of the context in which the decision is taken impact altruism a lot by changing the perceived norm (List, 2007; Bardsley, 2008), by providing excuses not to give (Dana et al., 2006) or by reallocating the property rights (Cherry et al., 2002) for instance. A premature conclusion could be that social preferences are very erratic by nature, which would cast doubt on their predictive power. An alternate conclusion it that social preferences are rather stable within individuals but the understanding of behavior driven by altruism, fairness concern or inequity aversion requires a thorough examination of the circumstances of the decisions. This examination in turn can provide interesting insights on the social and psychological motivations of individuals (Falk and Heckman, 2009).

Bibliography

Abeler, J., D. Nosenzo, and C. Raymond (2018). Preferences for truth-telling. *Econometrica*.

Akerlof, G. and J. Yellen (1988). Fairness and unemployment. *American Economic Review 78*(2), 44–49.

Akerlof, G. A. (1982). Labor contracts as partial gift exchange*. *The Quarterly Journal of Economics 97*(4), 543–569.

Akerlof, G. A. and R. E. Kranton (2000). Economics and identity. *The Quarterly Journal of Economics 115*(3), 715–753.

Akerlof, G. A. and J. L. Yellen (1990). The fair wage-effort hypothesis and unemployment. *The Quarterly Journal of Economics 105*(2), 255–283.

Andreoni, J. (1995). Cooperation in public-goods experiments: kindness or confusion? *The American Economic Review*, 891–904.

Andreoni, J. and B. D. Bernheim (2009). Social image and the 50-50 norm: A theoretical and experimental analysis of audience effects. *Econometrica 77*(5), 1607–1636.

Andreoni, J. and R. Petrie (2004). Public goods experiments without confidentiality: a glimpse into fund-raising. *Journal of Public Economics 88*(7), 1605 – 1623.

Andreoni, J., J. M. Rao, and H. Trachtman (2017). Avoiding the ask: A field experiment on altruism, empathy, and charitable giving. *Journal of Political Economy 125*(3), 625–653.

Ariely, D., A. Bracha, and S. Meier (2009, March). Doing good or doing well? image motivation and monetary incentives in behaving prosocially. *American Economic Review 99*(1), 544–55.

Arrow, K. (1973). The theory of discrimination. *Discrimination in labor markets 3*(10), 3–33.

Ashraf, N., C. F. Camerer, and G. Loewenstein (2005, September). Adam smith, behavioral economist. *Journal of Economic Perspectives 19*(3), 131–145.

Bardsley, N. (2008, Jun). Dictator game giving: altruism or artefact? *Experimental Economics 11*(2), 122–133.

Becker, G. S. (1974). A theory of social interactions. *Journal of Political Economy 82*(6), 1063–1093.

Bénabou, R. and J. Tirole (2006, December). Incentives and prosocial behavior. *American Economic Review 96*(5), 1652–1678.

Benjamin, D. J., J. J. Choi, and A. J. Strickland (2010). Social identity and preferences. *American Economic Review 100*(4), 1913–28.

Berg, J., J. Dickhaut, and K. McCabe (1995). Trust, reciprocity, and social history. *Games and Economic Behavior 10*(1), 122 – 142.

Berman, E. (2000). Sect, subsidy, and sacrifice: An economist's view of ultra-orthodox jews. *The Quarterly Journal of Economics 115*(3), 905–953.

Bernheim, B. D. (1994). A theory of conformity. *Journal of Political Economy 102*(5), 841–877.

Bolton, G. E. and A. Ockenfels (2000). Erc: A theory of equity, reciprocity, and competition. *American economic review 90*(1), 166–193.

Brandts, J. and N. Figueras (2003). An exploration of reputation formation in experimental games. *Journal of Economic Behavior & Organization 50*(1), 89 – 115.

Bursztyn, L. and R. Jensen (2015). How does peer pressure affect educational investments? *The Quarterly Journal of Economics 130*(3), 1329–1367.

Camerer, C. and K. Weigelt (1988). Experimental tests of a sequential equilibrium reputation model. *Econometrica 56*(1), 1–36.

Carvalho, J.-P. (2013). Veiling. *The Quarterly Journal of Economics 128*(1), 337–370.

Cettolin, E. and S. Suetens (2018). Return on trust is lower for immigrants. *The Economic Journal*.

Charness, G., D. Masclet, and M. C. Villeval (2014). The dark side of competition for status. *Management Science 60*(1), 38–55.

Charness, G. and M. Rabin (2002). Understanding social preferences with simple tests. *The Quarterly Journal of Economics 117*(3), 817–869.

Charness, G., L. Rigotti, and A. Rustichini (2007, September). Individual behavior and group membership. *American Economic Review 97*(4), 1340–1352.

Chen, R. and Y. Chen (2011, October). The potential of social identity for equilibrium selection. *American Economic Review 101*(6), 2562–89.

Chen, Y. and S. X. Li (2009, March). Group identity and social preferences. *American Economic Review 99*(1), 431–57.

Cherry, T. L., P. Frykblom, and J. F. Shogren (2002). Hardnose the dictator. *The American Economic Review 92*(4), 1218–1221.

Cohn, A., M. A. Maréchal, and T. Noll (2015). Bad boys: How criminal identity salience affects rule violation. *The Review of Economic Studies 82*(4), 1289–1308.

Dana, J., R. a. Weber, and J. X. Kuang (2006, sep). Exploiting moral wiggle room: experiments demonstrating an illusory preference for fairness. *Economic Theory 33*(1), 67–80.

Duesenberry, J. (1949). *Income: Saving, and the Theory of Consumer Behavior*. Harvard economic studies. HarvardUniversity Press.

Falk, A. and J. J. Heckman (2009). Lab experiments are a major source of knowledge in the social sciences. *science 326*(5952), 535–538.

Fehr, E., U. Fischbacher, and S. Gächter (2002). Strong reciprocity, human cooperation, and the enforcement of social norms. *Human nature 13*(1), 1–25.

Fehr, E., G. Kirchsteiger, and A. Riedl (1993). Does fairness prevent market clearing? an experimental investigation. *The Quarterly Journal of Economics 108*(2), 437–459.

Fehr, E. and K. M. Schmidt (1999). A theory of fairness, competition, and cooperation. *The quarterly journal of economics 114*(3), 817–868.

Fischbacher, U. and F. Föllmi-Heusi (2013). Lies in disguiseâ€"an experimental study on cheating. *Journal of the European Economic Association 11*(3), 525–547.

Forsythe, R., J. L. Horowitz, N. Savin, and M. Sefton (1994). Fairness in simple bargaining experiments. *Games and Economic Behavior 6*(3), 347 – 369.

Friedman, M. (1953). *Essays in Positive Economics*. A Phoenix book. Business economics. University of Chicago Press.

Gneezy, U. (2005). Deception: The role of consequences. *American Economic Review 95*(1), 384–394.

Gneezy, U. and A. Rustichini (2000). Pay enough or don't pay at all. *The Quarterly Journal of Economics 115*(3), 791–810.

Goeree, J. K., C. A. Holt, and S. K. Laury (2002). Private costs and public benefits: unraveling the effects of altruism and noisy behavior. *Journal of Public Economics 83*(2), 255 – 276.

Güth, W., R. Schmittberger, and B. Schwarze (1982). An experimental analysis of ultimatum bargaining. *Journal of Economic Behavior & Organization 3*(4), 367 – 388.

Harbaugh, W. T. (1998). The prestige motive for making charitable transfers. *The American Economic Review 88*(2), 277–282.

Hargreaves Heap, S. P. and D. J. Zizzo (2009, March). The value of groups. *American Economic Review 99*(1), 295–323.

Hedegaard, M. S. and J.-R. Tyran (2018, January). The price of prejudice. *American Economic Journal: Applied Economics 10*(1), 40–63.

Hoffman, E., K. McCabe, K. Shachat, and V. Smith (1994). Preferences, property rights, and anonymity in bargaining games. *Games and Economic behavior 7*(3), 346–380.

Huck, S. and G. K. Lünser (2010). Group reputations: An experimental foray. *Journal of Economic Behavior & Organization 73*(2), 153–157.

Huck, S., G. K. Lünser, and J.-R. Tyran (2012). Competition fosters trust. *Games and Economic Behavior 76*(1), 195–209.

Kahneman, D., J. L. Knetsch, and R. Thaler (1986). Fairness as a constraint on profit seeking: Entitlements in the market. *The American economic review*, 728–741.

Kahneman, D. and A. Tversky (1979). Prospect theory: An analysis of decision under risk. *Econometrica 47*(2), 263–291.

Kajackaite, A., U. Gneezy, and J. Sobel (2018). Lying aversion and the size of the lie. *The American Economic Review*.

Köszegi, B. and M. Rabin (2006). A model of reference-dependent preferences. *The Quarterly Journal of Economics 121*(4), 1133–1165.

Lane, T. (2016). Discrimination in the laboratory: A meta-analysis of economics experiments. *European Economic Review 90*, 375 – 402. Social identity and discrimination.

List, J. A. (2007). On the Interpretation of Giving in Dictator Games. *Journal of Political Economy 115*(3), 482–493.

Mill, J. S. (1874). *Essays on Some Unsettled Questions of Political Economy* (2 ed.). McMaster University Archive for the History of Economic Thought.

Montmarquette, C., J.-L. Rullière, M.-C. Villeval, and R. Zeiliger (2004). Redesigning teams and incentives in a merger: An experiment with managers and students. *Management Science 50*(10), 1379–1389.

Phelps, E. S. (1972). The statistical theory of racism and sexism. *The american economic review*, 659–661.

Riley, J. G. (2001). Silver signals: Twenty-five years of screening and signaling. *Journal of Economic literature 39*(2), 432–478.

Shariff, A. F. and A. Norenzayan (2007). God is watching you: Priming god concepts increases prosocial behavior in an anonymous economic game. *Psychological science 18*(9), 803–809.

Smith, A. (1759). *The Theory of Moral Sentiments*. McMaster University Archive for the History of Economic Thought.

Smith, V. L. (2012). Adam smith on humanomic behavior. *The Journal of Behavioral Finance & Economics 2*(1), 1–20.

Smith, V. L. (2013). Adam smith: From propriety and sentiments to property and wealth. *Forum for Social Economics 42*(4), 283–297.

Spence, M. (1973). Job market signaling. *The Quarterly Journal of Economics 87*(3), 355–374.

Tajfel, H. and J. C. Turner (1979). An integrative theory of intergroup conflict. *The social psychology of intergroup relations 33*(47), 74.

Tirole, J. (1996). A theory of collective reputations (with applications to the persistence of corruption and to firm quality). *The Review of Economic Studies 63*(1), 1–22.

Tversky, A. and D. Kahneman (1991). Loss aversion in riskless choice: A reference-dependent model. *The Quarterly Journal of Economics 106*(4), 1039–1061.

Wilson, R. (1985). Reputations in games and markets. *Game-theoretic models of bargaining*, 27–62.

Chapter 1

Does upward mobility harm trust? [1]

1.1 Introduction

Social mobility, defined as the movement from a given social level to a higher level, is considered as a cornerstone of modern democracies because it is a fundamental component of justice (Sen, 1980). Within organizations, upward mobility consists in internal promotions, which are a fundamental source of motivation (Prendergast, 1999; Lazear and Gibbs, 2014; Kuhn, 2017). Systematically hiring on the external labor market for filling higher positions might be met with resent by incumbents and reduce their willingness to cooperate.

However, social mobility may also have less positive side effects, in particular when it affects negatively the interpersonal relationships of upwardly mobile individuals with others. For example, females who achieve top positions in male dominated organizations face sometimes difficulties integrating the informal networks of leaders because these networks are usually composed of males (e.g., the “Old Boy Networks", Oakley, 2000; Goldin, 2014); when they succeed, they are not always willing to help other females to reach the same positions (this has been described as the “Queen-Bee effect" Ellemers et al., 2012). The upwardly mobile individuals are also sometimes rejected by

[1] Joint work with Marie Claire Villeval. R&R at the Journal of Economic Behavior and Organization.

left behinds in the process of social mobility. The socially mobile individuals may no longer be considered as in-groups by the left behinds, while not being treated yet by higher status individuals as one of them.

Our research objective is to test whether individual upward mobility affects trust in the presence of group identity. We are interested in situations in which each individual belongs to a well identified group (e.g. gender, ethnicity, or school) and also has a status (e.g. leadership or authority) that is correlated with his group. We study the specific experience of those who achieve a higher status than what their group predicted (e.g. female who achieve top positions in male dominated organizations) and how their promotion to a higher status affects their relations with ingroups and outgroups. Our premise is that upward mobility might mitigate group identity by increasing social distance between ingroups of different status and reducing social distance between outgroups of equal status. We focus on trust because of its importance in promoting efficient economic interactions (Arrow, 1972), especially within organizations, where interpersonal trust is a substitute to costly monitoring and a potential response to incomplete contracting (Fukuyama, 1995; LaPorta et al., 1997). Crucially to our study, social distance and identity are utterly important determinants of trust. Indeed, the empirical literature has shown that individuals consistently discriminate by trusting members from other groups less than people from their own group (Fershtman and Gneezy, 2001; Heap and Zizzo, 2009; Falk and Zehnder, 2013). We design a laboratory experiment which captures the main features of upward mobility and allows to measure its effect on trust and trustworthiness.

Our controlled laboratory experiment is based on a standard trust game (Berg et al., 1995) in which trustors and trustees were identified both by their affiliation to a group identity and by a status. The group identity was conferred by natural group identities: subjects were identified by their school, either the local business school or the local engineering school. These two very selective schools have a specific cultural identity and students from each school have a very strong feeling of belonging. In the first part

of the experiment, we reinforced these natural identities and in the second part, we used dictator games to test the existence of in-group favoritism.

Status was induced by asking subjects to answer to a math quiz in the third part of the experiment. Relative performances in that task were used to assign positions of "experts" and "agents" to the subjects that determined the task people had to do in the last part of the experiment. The experts were the subjects who gave the higher number of correct answers. This introduced a hierarchy between schools since the curricula in the engineering school are naturally more math-oriented: it was more likely for the students from this school to achieve the higher status of experts. A critical feature of our design is that the school affiliation predicts performance in the math quiz and thus, predicts the achieved status, but imperfectly. Indeed, we represent upward social mobility by assigning the higher status of expert also to the best performers from the business school: these subjects achieved the status of expert although their school curricula made it ex ante unlikely. This was controlled by inviting 60% of the subjects from the business school and by having 60% of experts in each session. Our design also disconnects status from payoffs: as experts and agents have different roles but receive the same payoff, this allows us to isolate the pure effect of status and identity on trust and trustworthiness, independently of wealth effects.

Then, in the fourth part of the experiment subjects played a trust game under the strategy method. They made decisions both as trustor and as trustee, being informed of the school and the status of expert or agent of their counterpart in all possible combinations. Comparing the transfers of the experts from the business school (the promoted) to the transfers of the agents from the business school is informative of the impact of social mobility on the propensity to trust others and to be trustworthy. Comparing the transfers directed at the experts from the business school to those directed at the agents from the business school informs us on the impact of upward mobility on the propensity

of others to trust and to be trustworthy to upwardly mobile subjects.

Our results show that social mobility reduces trust toward subjects from the same school who did not achieve the status of experts but also toward subjects from the other school. This cannot be explained by social preferences, risk aversion or beliefs about trustee's trustworthiness. Thanks to an additional condition in which we recruited subjects from the business school only, we explore the role of inter-group comparison in the effect of upward mobility on trust. In this additional condition, experts from the business school trusted as much as agents. Taken together, these results are not consistent with our hypothesis according to which upward mobility increases social distance between the mobile and ingroups and decreases social distance between the mobile and outgroups of the same status. Alternatively, upward mobility seems to single promoted individuals out, due to the singularity of being awarded the high status as a subjects from the business school.[2] We also found that social mobility does not affect trustworthiness: upwardly mobile individuals reciprocate as much as left behinds, irrespective of the trustor's identity. This confirms that the effect of upward mobility does not transit through social preferences. Finally, we show that subjects from either school transfer similar amounts as trustor and as trustee to upwardly mobile individuals and to left behinds. Thus, social mobility in itself does not attract a discriminatory behavior, possibly because by design social mobility does not result from self-selection on such traits as greed or competitiveness.

The remaining of the paper is organized as follows. Section 1.2 reviews briefly the related economic literature. Section 1.3 presents the experimental design, the procedures and our behavioral conjectures. Section 1.4 displays our results and section 1.5 discusses these results and concludes.

[2]Galeotti and Zizzo (2014) for instance show that singling-out an individual in a group can impact his behavior in subsequent trust games.

1.2 Related literature

Our study contributes to advance two main strands of the literature. First, it relates to the literature on the impact of social mobility on social preferences. In particular, this literature has shown that the prospect or experience of upward mobility weakens the support for redistribution (Piketty, 1995; Alesina and Ferrara, 2005; Acemoglu et al., 2017; Alesina et al., 2017). Our approach is different. This previous literature focuses mostly on inter-generational income mobility whereas we consider social mobility in terms of status and hold income equal across statuses. Moreover, it considers individual support for redistribution, while we study interpersonal trust. More directly connected to our research are the studies of Austen-Smith and Fryer (2005) and Fryer and Torelli (2010) that analyze the relationships between social mobility and social preferences in the context of the racial gap in educational achievements. Students who achieve more than what was expected, given their race, tend to be rejected by those who are lagging behind. Heap and Zizzo (2009) and Tsutsui and Zizzo (2013) report on experiments in which subjects are split into groups, play trust games between and within groups and can periodically trade group membership on a market place. In Tsutsui and Zizzo (2013), groups have different statuses and subjects can pay to quit a low-status group and join a high-status group, which reveals a demand for upward mobility. In these studies, the mechanism behind mobility is self-selection: in equilibrium, mobility concerns those who identify less ex-ante to their background group. Our design is different since subjects cannot choose to join the high-status group.

Second, our study contributes to the large literature on the impact of group identity on economic behavior, as pioneered in economics by Akerlof and Kranton (2000). Group identity leads to in-group favoritism (e.g. Charness et al., 2007; Chen and Li, 2009) and parochial exclusion (e.g., Bowles and Gintis, 2004), influences norm enforcement (e.g., Bernhard et al., 2006), and fosters cooperation (e.g., Goette et al., 2006) and coordination (e.g., Chen and Chen, 2011). In this literature group identity is usually uni-dimensional:

individuals are affiliated to a single group. In contrast, in our experiment identity is conferred by two dimensions that overlap partially: a natural group identity (school affiliation) and a lab-induced identity conferred by assigning a status of expert or agent to the subjects.

Very few papers have considered multi-dimensional group identities. Klor and Shayo (2010) induce status on top of natural group affiliations to test the impact of group identity on redistribution. Kranton et al. (2016) compare within-subjects the impact of natural identity (political affiliation) and arbitrary minimal group identity on allocation choices. They show a new type of heterogeneity of preferences between subjects who do not care about group identities and others who change their preferences according to the affiliation of their counterparts. Chen et al. (2014) play coordination and cooperation games with subjects from the same school with different ethnic identities. They show that priming school fosters cooperation and coordination, while priming ethnic identities is harmful. We differ in several respects from these papers. In our experiment, status and material payoffs are uncorrelated ex ante, our subjects are identified by both group affiliations when they interact, natural identity is predictive of status, and there is a notion of hierarchy between groups. Hong et al. (2016) also introduce a hierarchy between groups by introducing both horizontal group identities, determined randomly or by preferences, and vertical group identities, determined by performance or luck. In contrast to them, we use natural identities, we keep income constant across groups, and we do not study redistributive allocations but interpersonal trust.

1.3 Design, Procedures and Conjectures

In this section, we first detail the experimental design (1.3.1). Then, we present the procedures (2.4). Last, we introduce our conceptual framework and behavioral conjectures (3)

1.3.1 Experimental Design

Before our subjects played trust games, we first reinforced natural group identities, then we measured the subjects' social preferences and we induced status. Figure A1.1 in A1 summarizes the timeline of the experiment.

Natural Group Identities

Our experiment uses natural group identities conferred by school membership. We recruited subjects from the local engineering school (Ecole Centrale de Lyon) and the local business school (Ecole de Management de Lyon). In the French higher education system, these schools are defined as "Grandes Ecoles", considered as providing education for elites, independently from Universities. Grandes Ecoles are relatively small in size (generally less than 4000 students, compared to more than 20 000 for universities) and they have very high educational requirements. The selection of students at entry is based on very competitive exams prepared intensively during two years. Each Grande Ecole has its own culture and traditions and belonging to such schools generates a strong group identity. The two schools have different curricula, but are equally selective.[3] In the instructions, it was made clear that the session was exclusively composed of students from these two schools.

We reinforced natural group identities by several means. Upon arrival, each computer screen displayed the logo of the subject's school. Then, the first part of the experiment consisted of a quiz about one's school. Subjects had four minutes to answer to six questions relative to their own school, including its year of creation, the expected wage after graduation, or famous alumni (see A3). Before submitting their individ-

[3]The high school exit exam grades of students from these schools support this claim. According to the french website "l'Étudiant" (hosting a database used by students in their educational choices), the average grade of student from both these schools was around 17/20. These are high grades since only 13% of the candidates have grade higher than 16/20. The data are from https://www.letudiant.fr/palmares/palmares-des-grandes-ecoles-de-commerce/em-lyon.html, https://www.letudiant.fr/palmares/palmares-des-ecoles-d-ingenieurs/centrale-lyon.html and http://www.education.gouv.fr/cid55597/resultats-definitifs-de-la-session-2017-du-baccalaureat-79-d-une-generation-est-titulaire-du-baccalaureat.html respectively (in French).

ual answers, subjects could communicate with their schoolmates in the session via an anonymous chatbox and discuss about the answers to the quiz. This procedure aims at reinforcing group identity (as in Chen and Li, 2009). The quiz was incentivized both at the individual and at the school level. Each correct answer yielded 1 Experimental Currency Unit (ECU, with 1 ECU = €0.2) to the subject. Moreover, each subject from the school who performed the best at the quiz earned an extra 5 ECU. This collective bonus aimed at activating both cooperation between schoolmates and competition between schools. No feedback about the absolute or the relative performance was provided to the subjects before the end of the session, to avoid creating wealth effects.

Social Preferences

In the second part, we elicited the subjects' social preferences toward in-groups from the same school and toward out-groups from the other school, using a dictator game. Subjects received a 10 ECU endowment and they had to decide how much to transfer to another subject, conditional on whether this subject is from the business school or from the engineering school. Subjects were informed that they would be randomly matched at the end of the session with a subject from each school. For one match (randomly drawn), their decision as dictator would be implemented, and for the other match, they would receive the transfer decided by the other subject. Behavior in these games is used to check whether school identity generates in-group favoritism. Moreover, social preferences are used as a control variable when analyzing trust and trustworthiness.

Status: Agents and Experts

Math Quiz In the third part we introduced status by means of two positions that were assigned to the subjets based on their relative performance at a mathematical quiz: "Agents" and "Experts". It is usually admitted that status is conferred by performance feedback, ranking, and public recognition.

At the beginning of the third part, the program randomly formed anonymous groups of five subjects, with three subjects from the business school and two from the engineering school; this was made common knowledge. Subjects had 15 minutes to solve 17 problems. The same problems were displayed in the same order on all the subjects' screens. For each problem, four possible solutions were proposed and subjects had to pick one (see A4). The difficulty of the problems increased gradually to maximize the variance of performance. Before performing this task, subjects were informed that their performance would impact their role in the remaining of the experiment, but with no more detail at this stage. The objective was to avoid self-selection into roles on traits such as preference for status, power or competitiveness. Once the 15 minutes had elapsed, subjects were informed that their performance would be used to award the position of Expert to the three subjects who performed the best in the quiz among the five players, while the position of Agent was assigned to the two remaining subjects. We chose to award status within each sub-sample of five subjects rather than at the session level for two reasons. First, because in each group of five there are two subjects from the engineering school for three positions of experts; this ensures that at least one subject from the business school will have the status of expert. Second, it provides some randomness in the attribution of status. Indeed, if we had awarded status at the session level, the subjects who performed the most at the session level would always have achieved the high status and it would have been hard to disentangle the effect of social mobility from that of abilities on behavior in the subsequent trust game.

The math quiz was incentivized. At the end of the session, the program selected randomly one subject in each sub-sample of five, and for each correct answer of the selected subject, each of the five subjects earned 1 ECU in addition to a fixed payoff of 5 ECU. This payment scheme ensures that each subject's expected payoff is equal within the sub-sample of five and that payment is kept independent from the position of expert or agent. Using instead an individual incentive scheme would have introduced payoff inequality between players and it would have been impossible to disentangle the effect

of inequality from the effect of status in the trust game.[4]

Using a math quiz to assign positions and introducing the notions of experts and agents aimed at generating a hierarchy in status associated with these positions. First, subjects from these schools are likely to acknowledge that math induces status because, in their studies, being good at math allowed them to enter famous schools. Second, curricula at the engineering school are more math oriented, then subjects from this school are likely to perform better on average than those from the business school, and then achieve the position of experts more often.[5] Finally, the notions of experts and agents reinforce the feeling of a hierarchy in status independent from income.

The math quiz also introduced an implicit hierarchy between schools, because subjects from the engineering school typically achieved the high status while those from the business school typically did not. This implicit hierarchy is found in real-life situations: promotion often implies that a member of a group that is low in the implicit hierarchy reaches the typical status of a group higher in the implicit hierarchy.[6] This might in turn reinforce the salience of upward mobility, since promoted subjects reach the status typical of subjects from the other school.

At the end of the quiz, subjects received the instructions for the rest of the experiment, namely an expertise task and the trust game. The expertise task was performed after playing the trust game to avoid contaminating behavior in the trust game, but subjects learned their status at the beginning of this part.

[4]Since we wanted to incentivize the math quiz in order to be sure that subjects provide efforts, we had to choose between an individual incentives scheme which introduces inequality, and a collective incentives scheme which might reinforce solidarity or generate reciprocity. We deemed the second less risky, because it is in subjects' best selfish interest to provide effort in the math quiz and such extrinsic motivations have been shown to mitigate reciprocity (see e.g. Stanca et al., 2009).

[5]Studying the selection technology of the French "Grandes Ecoles" Menger and Marchika (2014) shows that the schools that are ranked higher in their specialty put a greater coefficient on mathematics at the entrance exam. Math grades account for slightly more than one third of the total grades at the entrance exam in top engineering schools, and for around one fourth at the entrance exam in top business schools.

[6]One could for instance think of a graduate from a regular college, who ends up in the board of a large firm, the other board members being essentially from Ivy league universities.

Expertise Task This task was designed to reinforce the statuses of experts and agents when the task was explained to the subjects, without generating payoff inequality. The agents had to answer to three more multiple choice math questions of similar difficulty, displayed at once on their screen. Then, being informed of the distribution of the agents' answers at the session level, the experts had to vote to select the question to be used to determine payoffs. The question that received the most votes determined a payoff for all the subjects in the session, regardless of their school or position. This payoff was 10 times the mean rate of correct answers given by the agents in response to this question (e.g., if the rate was 0.8, each subject in the session earned 8 ECU). The position of expert is expected to provide higher status since the experts have to evaluate the work of the agents who execute the task, and take responsibility for a choice that determines the payoff of everyone.

A potential concern with instructing the expertise task before the trust game is that it could introduce reciprocity or reinforce solidarity between agents and experts in the trust game.[7] In order to asses the severity of this concern, we ran three sessions (45 subjects) of a robustness condition in which information about the expertise task given before the trust game is minimal: we only inform subjects that there will be a task in which experts and agents will have different role after the trust game and that payoffs will be equal across status, but we do not inform them of the precise content of the task or how payoffs will be determined. In this condition, solidarity or reciprocity concerns cannot stem from anticipation about the content of the expertise task. The data from this robustness treatment is statistically indistinguishable from the data from the main condition (See A5), so that we pool the data from the main and the robustness conditions in our analysis.

[7] We thank an anonymous referee for raising this potential concern.

Trust and Trustworthiness

We are chiefly interested in behavior in the trust game played in the fourth part, conditional on school, status and social mobility. In this game, both the first and the second movers receive a 10 ECU endowment. The first mover (the trustor, hereafter) has to choose an integer amount M between 0 and 10 ECU, inclusive, that will be sent to the second mover (the trustee, hereafter) and deducted from his payoff. This amount is multiplied by three. The trustee receives $3M$ and has to choose an integer amount $R \leq 3M$ to send back to the trustor. The amount sent by the trustor is a standard measure of trust, while the amount returned by the trustee is a measure of trustworthiness.

In each group of five, each subject first made four decisions as a trustor, interacting successively with each of the other four players in the role of trustees. Subjects had also to guess the proportion of the amount received each trustee would send back to them; they were rewarded for accuracy, as explained below. Finally, each subject had to make four decisions as a trustee, interacting successively with each of the four other players in the role of trustors. We used the strategy method: the trustees had to decide how much to return to the trustors for each of the ten possible transfers. Using the strategy method allows us to collect more data for each subject and to define a more accurate profile of the trustees' preferences. For each decision as trustor and as trustee, subjects were informed about the school and the status of their counterpart (expert or agent). At this moment, they could observe whether one or more players from the business school achieved the status of experts.

At the end of the experiment, the computer program randomly selected one decision in each role for each subject and the sum of payoffs in these two decisions determined the payoff for the trust game.[8] The program also randomly selected for each individual one guess made in the role of a trustor. If this guess was equal to the actual proportion

[8]The software generated two random numbers for each subject. These numbers were used to generate two rankings of subjects within each group of five. The first ranking was used to determine which decision was used to compute payoffs as trustor and the second ranking was used to determine which decision was

returned by the trustee more or less 5 percent, the subject earned an extra 5 ECU.

Finally, at the end of the session subjects had to answer to several questions about socio-demographic characteristics, risk attitudes (using the procedure of Dohmen et al., 2011), and perceptions about the experiment. In particular, they had report their opinion about the statement *"It is very important to have good math skills"* on a five point likert-scale. No subject strongly disagreed with the statement and 128 subjects out of 145 agreed to some extent.[9] They also had to report which school they believe have the best students in math. Answers strongly support that math induces a hierarchy between schools, since all the subjects from the engineering school and 9 out of 10 subjects from the business school reported that the students from the engineering school were the best at math.

The total payoff in the experiment was the sum of the payoffs made in the quiz about schools, in the dictator game, in the math quiz, in the trust game and in the expertise task.

1.3.2 Procedures

The experiment was developed using z-Tree (Fischbacher, 2007). All sessions were conducted at GATE-LAB, Lyon, France. We ran nine sessions (six in the main condition, three in the robustness condition) with 10 to 20 subjects each that were recruited using Hroot (Bock et al., 2014). In total, 145 subjects took part in the experiment: 87 from the Business school (Ecole de Management de Lyon) and 58 from the Engineering school (Ecole Centrale de Lyon). 41% of the subjects are females (36% of the subjects from the engineering school and 44% of those from the business school; Fisher's exact test, $p = 0.39$). The average age is 21.23 years (21.58 for the subjects from the engineering

used to compute payoffs as trustee. In both cases, the decision used for payment was the one in which the counterpart was next in the corresponding ranking.

[9]Subjects acknowledge the importance of math skills irrespective of their school or status. The distribution of opinions between "Strongly disagree", "Disagree", "Somewhat agree", "Agree", "Strongly Agree" differs neither across schools ($\chi^2(3) = 0.548$, $p = 0.908$), nor across statuses ($\chi^2(3) = 2.212$, $p = 0.529$).

school and 21 for those from the business school; Mann-Whitney test, $p = 0.038$).

Upon arrival, subjects drew a tag from one of two opaque bags (one for each school) assigning them to a cubicle. The instructions for each part were distributed and read aloud by the same experimenter after completion of the previous part, except the instructions for the trust game and the expertise task that were distributed together in the main condition (see A1). Before playing the trust game, subjects had to fill out a comprehension questionnaire. Questions were answered in private.

The average duration of sessions was 80 minutes. The average payoff was €14.49 (Min: €9.10, Max: €21.60, standard deviation, S.D. hereafter: 2.30). Payments were made in cash, in a separate room and in private.

1.3.3 Conceptual Framework and Conjectures

In the trust game, without social preferences there is a unique subgame perfect Nash equilibrium. The trustee has no incentive to return any amount to the trustor and the trustor anticipates this. Hence, both the trustor and the trustee send nothing. Empirically, however, behavior usually deviates from the SPNE, as trustors and trustees typically send non trivial amounts.[10] Moreover, various studies have shown that trust and trustworthiness are generally higher with ingroups than with outgroups. This is true for artificial groups (e.g., Buchan et al., 2006) or natural groups (e.g., Etang et al., 2011).[11] Social distance, defined as "the perceived distance between individuals and groups" (Kazdin, 2000; Dufwenberg and Muren, 2006), reduces trust between individuals.

In our experiment, each subject is characterized both by the natural group identity conferred by his school and by the status conferred by the position of agent or expert,

[10] In a meta-analysis, Johnson and Mislin (2011) found that trustors send almost half of their endowment and trustee return about 35% of the amount they received.

[11] Behavior in the trust game depends on social preferences and on the expectations about the counterpart's social preferences. For instance, Chen and Li (2009) showed that people are kinder to in-groups, and Goette et al. (2006) found that people expect more kindness from in-groups than from out-groups.

conditional on his relative performance at the math quiz. Hence, these two dimensions allow us to define four categories. For the following two categories, the status is consistent with the natural group identity. The “Steady low" are the subjects from the business school who have been assigned the status of agents, and the “Steady high” are the subjects from the engineering school who are experts. The two other categories capture two opposite forms of social mobility: promotion and demotion. We identify as “Promoted” the subjects from the business school who have achieved the status of expert. Indeed, due to their education, they were less likely ex-ante to reach this status compared to the students from the engineering school. We assume that assigning them the position of experts captures the idea of a symbolic promotion to the higher status position. Last, “Demoted" are the subjects from the engineering school who have unexpectedly achieved the status of agents.[12] These denominations were not mentioned in the instructions. Table 1.1 characterizes the four categories obtained by the combination of natural groups and the statuses.

Status / Natural group	Agent	Expert
Engineering school	Demoted	Steady-High
Business school	Steady-Low	Promoted

Table 1.1 – Four categories of subjects in our experiment

Our experiment basically revolves around assessing the interaction effects of a natural group identity and a status. At first sight, the natural group identity seem stronger, but status has been shown to have strong behavioral impacts even with no monetary consequence and no value beyond the time frame of the experiment (e.g. Charness et al., 2013). This pattern is explained by the fact that people use status to positively distinguish themselves from others and enhance their self-esteem (Tajfel and Turner, 1979). Thus, they consider status an important part of their identity especially if it fulfills this objective. We primarily expect our lab-induced status to mitigate the ingroup bias

[12]Our experiment was designed to study upward social mobility. We only have fourteen demoted individuals. Thus, we leave the study of the impact of downward mobility for further research.

generated by natural identities and then increase social distance between experts and agents conditional on ingroup / outgroup. In other words, we expect less ingroup bias between subjects from the same school but with different status than between subjects from the same school and with the same status. Symmetrically, we expect less ingroup bias between subjects from different schools but with the same status than between subjects from different schools and with different statuses. The previous discussion gives the following conjectures.

Conjecture 1: Compared to steady low subjects, the promoted send less to steady low subjects, both as trustors and as trustees.

Conjecture 2: Compared to steady low subjects, the promoted send more to steady high subjects, both as trustors and as trustees.

Conjecture 3: Compared to steady low subjects, the promoted receive higher amounts from steady high, both as trustors and as trustees.

Conjecture 4: Compared to steady low, the promoted receive lower amounts from steady low, both as trustors and as trustees.

1.4 Results

We first present summary statistics about status, social preferences, trust and trustworthiness (1.4.1). Then, we expose our results on the behavior of promoted subjects (1.4.2) and

on behavior towards promoted subjects (1.4.3). Finally, we discuss the role of intergroup comparison thanks to the data from an additional treatment (1.4.4).

1.4.1 Summary statistics

Status Status was assigned via the math quiz. The average number of correct answers in the math quiz is 5.83 (Min: 1, Max: 13, S.D.: 2.48). As expected, this number is significantly higher for the subjects from the engineering school (7.15) than for those from the business school (4.95) (Mann-Whitney, M-W hereafter: $p < 0.001$).[13] The lowest performance for an expert from the engineering school (the business school, respectively) is 3 (4, resp.), and the best performance is 13 (9, resp.). On average, the number of correct answers is 7.19 for the experts and 3.79 for the agents (M-W: $p < 0.001$).

By design, 87 subjects achieved the status of experts. The probability to reach this status is 0.76 for a subject from the engineering school and 0.49 for a subject from the business school (Fisher's exact test: $p = 0.002$). Among the 58 subjects from the engineering school, 44 became experts (steady-high) and 14 failed to do so (demoted). Among the 87 subjects from the business school, 43 became experts (promoted) and 44 did not (steady-low). In thirteen groups out of 29, there is more than one promoted subject (one with three promoted subjects and twelve with two promoted subjects).[14]

Social preferences We elicited social preferences by letting subjects play dictator games with a receiver either from the same school or from the other school. On average, dictators transfer 2.58 to a receiver from the same school and 1.56 to a receiver from the other school (Wilcoxon signed-rank test, W hereafter: $p < 0.001$). In-group favoritism is observed in both schools (these numbers are respectively 2.63 and 1.55 for the dictators

[13] All the tests are two-sided and take each individual as one independent observation.

[14] Note that the distributions of performance of experts and agents from the business school overlap. Figure A6.1 in A6 depicts the distribution of performances in the math quiz of the subjects from the business school, by status. This is due to the fact that status was awarded within the sub-samples of five subjects and not at the session level. The advantage is that performance at the task is not a perfect predictor of status, which is interesting for identification purposes.

from the business school (W: $p < 0.001$), and 2.51 and 1.58 for the dictators from the engineering school (W: $p < 0.001$). The transfers to a receiver from the same school do not differ across schools (M-W : $p = 0.933$). Transfers do not differ either when the receiver is from the other school (M-W: $p = 0.836$). Finally, we checked that promoted and steady-low do not differ before status is awarded in terms of social preferences. Their transfers do not differ, regardless of whether the receiver is from the same school (M-W : $p = 0.623$), or from the other school ($p = 0.951$).

Trust and Trustworthiness Table 1.2 displays summary statistics about trust, trustworthiness and beliefs, for the whole sample of subjects, by school and by status. The average transfer by trustors is 3.37 ECU (S.D.: 2.8). Transfers do not differ across schools (3.15 for the engineering school and 3.51 for the business school. M-W: $p = 0.491$).[15] Regarding beliefs, trustors guess that the trustees will return on average 24.69% of what they received. Beliefs do not differ across schools (24.1% for the engineering school and 25% for the business school. M-W: $p = 0.762$). Trustees return on average 19.7% of what they received (S.D.: 19.23). Trustworthiness does not differ across schools either (19.87% for the engineering school and 19.64% for the business school. M-W: $p = 0.918$). Table 1.2 also indicates that there is no significant difference in trust, beliefs and trustworthiness between agents and experts from either school.

Let us now consider in-group favoritism. Pooling both schools, the average amount sent to a trustee from the same school is 3.62 ECU and the average amount sent to a trustee from the other school is 3.24 ECU (W: $p = 0.012$). Considering each school separately shows evidence of an asymmetrical in-group bias in trust: trustors from the engineering school send more to trustees from the engineering school than to out-groups (3.63 ECU *vs.* 3, W: $p = 0.006$), whereas trustors from the business school send similar amounts to in-group and out-group trustees (3.61 *vs.* 3.40, W: $p = 0.344$). The same asymmetry is found in beliefs. Trustors from the engineering school expect a higher

[15]For this non-parametric test we average for each individual all his decisions as a trustor, so that each subject gives one independent observation.

	Aggregate (1)	Agents (2)	Experts (3)	$p-value$ (2)-(3)
Trust (mean amount sent)				
Business school	3.51 (2.85)	3.92 (2.75)	3.08 (2.91)	0.102
Engineering school	3.15 (2.89)	2.85 (2.55)	3.25 (2.84)	0.847
$p-value$	0.491	0.236	0.685	-
Trustworthiness (mean % of the tripled amount received that is returned)				
Business school	17.60 (19.52)	20.21 (17.41)	19.07 (21.09)	0.528
Engineering school	15.89 (18.65)	13.98 (14.5)	21.74 (20.55)	0.243
$p-value$	0.918	0.269	0.455	-
Beliefs (mean % of the tripled amount expected in return)				
Business school	25.08(21.75)	24.94 (22.54)	25.23 (21.17)	0.854
Engineering school	24.10 (21.1)	22.60 (23.21)	24.57 (20.65)	0.576
$p-value$	0.762	0.553	0.911	-

Notes: Standard deviations are in parentheses. The $p-values$ in the last column are from M-W tests comparing agents and experts. The $p-values$ in lines are from M-W tests comparing subjects from the two schools. The average of decisions of each subject gives one independent observation.

Table 1.2 – Summary statistics on trust, trustworthiness and beliefs, by school and status

return from their in-group trustees than from their out-group trustees (28.15 ECU *vs.* 22.75, W: $p = 0.014$), whereas trustors from the business school expect similar returns from both types of trustees (24.54 ECU *vs.* 25.63, W: $p = 0.966$).[16]

In contrast, the in-group favoritism in trustworthiness is symmetric across schools: trustees from both schools return more to trustors from the same school than to trustors from the other school (21.7 ECU *vs.* 18.3, W: $p < 0.001$).

1.4.2 Behavior of the Promoted

In this section, we focus on the behavior of the promoted in the trust game. We introduce our first result:

Result 1: Compared to steady-low, the promoted trust less. This is true for both steady-low and steady-high trustees.

[16]Note that this asymmetry is not problematic for our analysis since we focus on the comparison between the promoted and the steady-low. The specific structure of the in-group bias is irrelevant from that respect.

Result 1 supports Conjecture 1, but not Conjecture 2.

Support for Result 1: Figure 1.1 plots the average amount sent by the trustors from the business school, depending on whether the trustor is promoted or not. The left panel focuses on the matches with trustees from the business school and the right panel on the matches with trustees from the engineering school. On average, the steady-low trustors send 4.02 ECU (S.D.: 2.82) to trustees from the business school, while the promoted trustors send 3.19 ECU (S.D.: 3.25, M-W: $p = 0.085$). On average, the steady-low send 3.82 ECU to trustees from the engineering school (S.D.: 3.09), and the promoted send 2.97 ECU (S.D.: 3.07, M-W: $p = 0.132$).

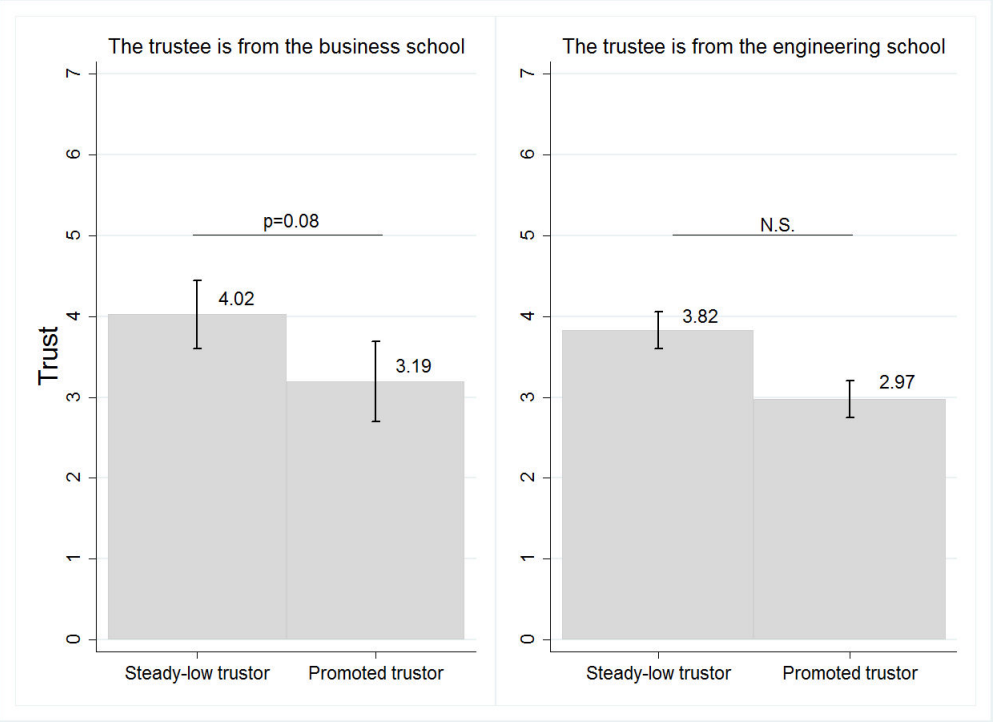


Figure 1.1 – Average amount sent by trustors from the business school, by trustor's status and trustee's school.

Note: The vertical lines represent standard errors.

The previous comparisons do not control for beliefs and individual characteristics. Table 1.3 reports marginal effects from Tobit regressions explaining the trusting decisions of subjects from the business school. We use Tobit models since data are censored, and robust standard errors are clustered at the individual level since each trustor makes several decisions.

Model (1) pools the observations of all possible matches. The independent variables include dummy variables for each category of match, taking the steady-low trustors matched with a steady-low trustee as the reference category. They include the amount the trustor expects to receive in return, social preferences toward in-groups and out-groups, risk attitudes, and gender. We include the performance in the math quiz as a proxy for higher intellectual abilities, because previous studies have suggested that more intelligent people trust more (e.g., Burks et al., 2009; Corgnet et al., 2016; Falk et al., 2018). We add a dummy variable indicating that the trustor somewhat or strongly disagrees with the statement about the importance of math skills, as it could capture a negative perception of the legitimacy of the selection of experts. Finally, session fixed effects are included.

In model (2), we restrict the observations to the cases in which the trustor is matched with a trustee who is also from the business school. We drop from the independent variables the categories of matches involving trustees from the engineering school. In model (3), we restrict the observations to the cases in which the trustor is matched with a trustee from the engineering school. In this model, the reference category is the steady-low trustors matched with a steady-high trustee.

Overall, these regressions show that promotion reduces trust, toward both steady-low and steady-high trustees.[17] Indeed, promoted trustors trust significantly less (at the

[17] Additional Tobit regressions, reported in Table A7.1 in A7, in which the dependent variable is the trustors' belief about the percentage returned by the trustee indicate that the effect of promotion on trust is not driven by differences in beliefs. Indeed, the beliefs of the promoted individuals do not differ significantly

Table 1.3 – Trust by subjects from the business school.

	Tobit models					
	(1)		(2)		(3)	
	Trust		Trust		Trust	
Beliefs	0.054****	(0.014)	0.052***	(0.017)	0.056****	(0.015)
Prom. trustor and Prom. trustee	-1.239	(1.121)	-1.239	(1.114)	-	-
Prom. trustor and St. low trustee	-2.013**	(0.922)	-2.002**	(0.941)	-	-
Prom. trustor and St. high trustee	-2.477***	(0.870)	-	-	-2.000**	(0.907)
Prom. trustor and Dem. trustee	-1.811*	(1.087)	-	-	-1.315	(1.170)
St. low trustor and Prom. trustee	-0.476	(0.531)	-0.510	(0.502)	-	-
St. low trustor and St. high. trustee	-0.461	(0.502)	-	-	-	-
St. low trustor and Dem. trustee	-1.689*	(0.894)	-	-	-1.159	(0.808)
Transfer in the DG, same school	0.287	(0.183)	0.356*	(0.191)	0.223	(0.190)
Transfer in the DG, other school	0.368	(0.225)	0.239	(0.228)	0.491**	(0.233)
Female	-0.377	(0.718)	-0.665	(0.747)	-0.0713	(0.774)
Risk attitude[a]	-0.057	(0.145)	0.080	(0.158)	-0.206	(0.163)
Math quiz performance	0.039	(0.246)	0.007	(0.259)	0.068	(0.282)
Perception Math[b]	-0.056	(0.843)	-0.227	(0.863)	0.097	(0.874)
Session F.E.	YES		YES		YES	
N	348		174		174	
Pseudo R^2	0.105		0.114		0.108	
Log pseudolikelihood	-734.264		-368.997		-359.224	
F	4.22		4.38		4.41	
$p > F$	< 0.001		< 0.001		< 0.001	

Notes: Marginal effects are reported. Robust standard errors clustered at the individual level are in parentheses. ** $p < 0.05$, *** $p < 0.01$. (1) Trustees from the business school and from the engineering school pooled, (2) Trustees from the business school only, (3) Trustees from the engineering school only.

The reference category in models (1) and (2) corresponds to steady-low trustors matched with a steady-low trustee; in model (3), it is a steady-low trustor matched with an steady-high trustee. Prom. for promoted (business school), St. low for steady-low (business school), steady-high for steady-high (engineering school), and Dem. for demoted (engineering school). (a) For risk attitudes, a higher number indicates less risk aversion. (b) Dummy indicating that the trustor chose "strongly disagree" or "disagree" about the statement *"It is important to have good mathematical skills"*.

5% level) when matched with a steady-low trustee, compared to steady-low trustors (see models (1) and (2)). Model (3) indicates that they also trust less the steady-high trustees from the engineering school than the steady-low trustors do. We also observe that trustors send more when they expect a higher return, which replicates previous findings in the literature. Most of the other independent variables are not significant, except that more generous subjects in the dictator game with an in-group (resp. out-group) trust also more when matched with an in-group (resp. out-group).[18]

Result 2: Promotion does not impact trustworthiness, regardless of the trustor's identity.

Support for Result 2: Compared to steady-low trustees, the promoted trustees do not return less to steady-low trustors and more to steady-high trustors. This result rejects Conjectures 1 and 2. Figure 1.2 displays the average percentage of the transfer received that is returned by trustees from the business school, depending on whether the trustee is promoted or not. The left panel focuses on the matches with trustors from the business school, the right panel on the matches with trustors from the engineering school. On average, the steady-low trustees return 21.89% of the amount received to trustors from the business school (S.D.: 18.13) and the promoted trustees return 20.99% (S.D.: 22.26, M-W: $p = 0.958$). On average, the steady-low trustees return 18.52% (S.D.: 17.87) to trustors from the engineering school and the promoted trustees return 17.15% (S.D.: 20.79, M-W: $p = 0.329$). None of these differences is significant. The pattern is similar if we only consider the returns to steady-high trustors: on average, the steady-low trustees return 22.36% to the steady-high trustees (S.D.: 22.33), while the promoted trustees return 17.77% (S.D.: 17.61) (MW: $p = 0.465$).

from those of the steady-low. Differences in risk attitudes cannot explain either the differences in trust. The mean measure of risk attitudes is 5.11 (S.D.: 2.29) for the steady-low and 5.30 for the promoted individuals (S.D.: 2.14); the difference is not significant (MW, $p = 0.66$).

[18] In Table A6.1 of A6, we present a robustness check on the sub-sample of subjects whose score in the math quiz was 3 or more, i.e. subjects whose score was enough to make them "promoted" in a least one group. We qualitatively replicate the results presented here, which shows that math abilities do not drive this result.

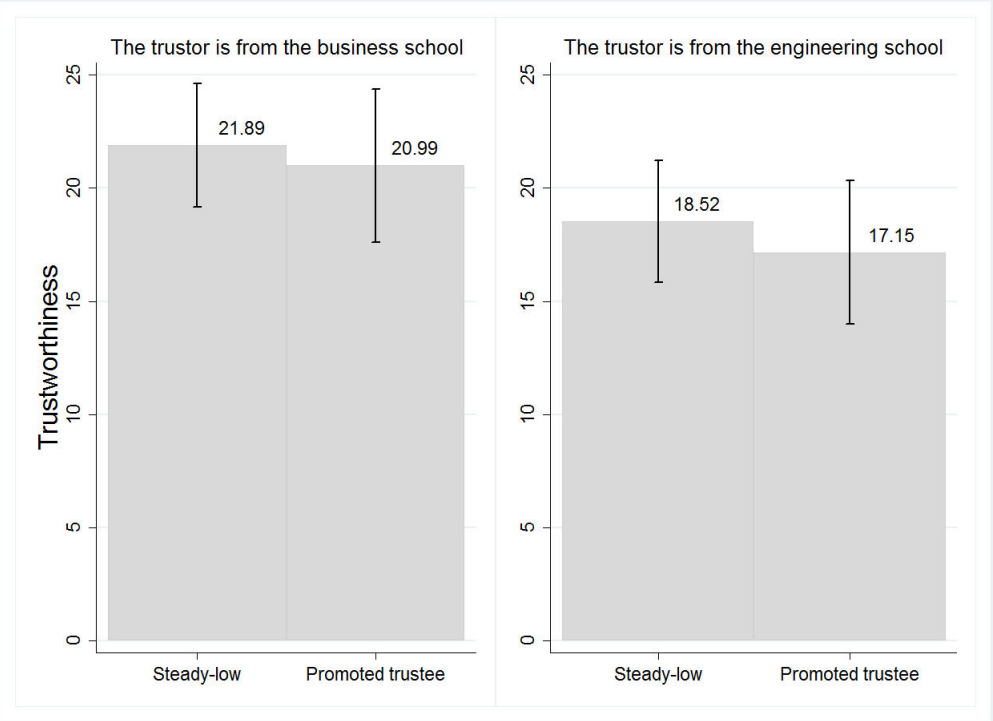


Figure 1.2 – Average percentage of the transfer that is returned by trustees from the business school, by trustee's status and trustor's school.

In A7, Table A7.2 presents the results of regressions that estimate Tobit models in which the dependent variable is the percentage of the amount received that is returned by the trustee, for trustees from the business school. For each interaction, we average the percentages returned for the different possible amounts received. The independent variables are the same as in the regressions on trust, except that we replaced the beliefs about returns by the amount transferred by the subject when acting as a trustor with the same counterpart. These regressions show that even when controlling for covariates, the trustee's promotion does not impact his trustworthiness.

1.4.3 Attitudes Toward The Promoted

In this section we compare the behavior of subjects when they are matched with promoted individuals or with steady-low from the business school. This analysis leads to the following result:

Result 3: Subjects from either school do not condition their decisions on whether their counterpart is promoted. This is observed for both trustors and trustees.

Conjectures 3 and 4 predicted that promoted individuals would receive more trust and trustworthiness from steady-high, and less trust and trustworthiness from steady-low. We find no evidence supporting these conjectures.

Support for Result 3: On average, promoted trustees receive 3.54 ECU (S.D. : 3.10), while steady-lows receive 3.27 ECU (S.D.: 2.98, W: $p = 0.177$) from non promoted trustors. Figure 1.3 displays the average trust directed to trustees from the business school by non-promoted trustors and separates promoted and steady-low trustees. The left panel focuses on the matches with trustors from the business school. On average, the promoted trustees receive 4.06 ECU from trustors from the business school (S.D.: 3.04) and the steady low trustees receive 3.96 ECU (S.D.: 3.29, W: $p = 0.712$). The right panel focuses on the matches with trustors from the engineering school. On average, the promoted trustees receive 2.88 ECU (S.D.: 2.74) and the steady-lows receive 3.14 ECU (S.D.: 3.11, W: $p = 0.156$). The conclusion is similar if we restrict the analysis to the steady-high trustors: on average, they send 2.96 ECU (S.D.: 2.98) to steady low trustees and 2.96 (S.D.: 2.98) to promoted trustees (W: $p = 0.231$).

The same holds for trustworthiness. Figure 1.4 represents the average percentage of the amount received from trustors of the business school that is returned by the non-promoted trustees. While, on average, steady-low trustors receive back 19.77% of the tripled amount sent (S.D.: 18.64), promoted trustors receive 20.54% of this amount (S.D.: 19.02, W: $p = 0.553$). The left panel focuses on the matches with trustees from the business school. On average, the steady-low trustors receive 20.93% (S.D.: 18.07) from trustees from the business school, and the promoted trustors receive 21.9% (S.D.: 18.15, W: $p = 0.206$). The right panel focuses on the matches with trustees from the engineering school. On average, the steady-low trustors receive 19.10% (S.D.: 19.09) from trustees from the engineering school, and the promoted trustors receive 19.50%

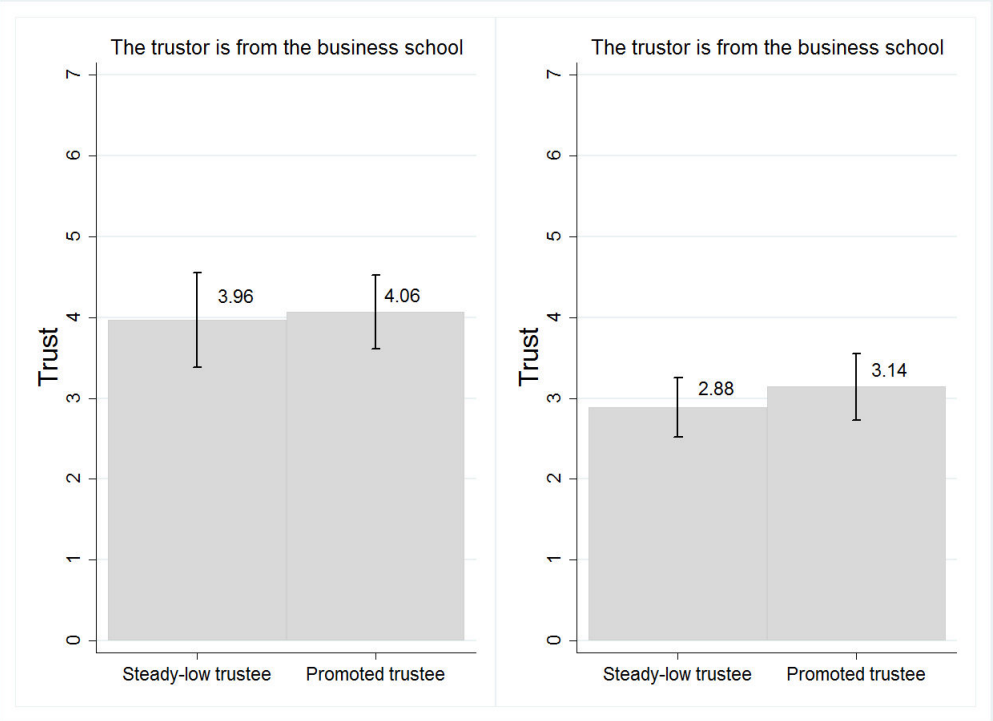


Figure 1.3 – Average transfers to trustees from the business school, by trustee's status and trustor's school.

(S.D.: 19.74, W: $p = 0.847$). Again, none of the pairwise comparisons reaches standard levels of significance and the conclusion remains the same if we consider the steady-high trustees who, on average, return 20.29% (S.D.: 22.67) to steady-low trustors and 19.79% (S.D.: 21.03) to promoted trustors (W: p=0.275)

In A7, Table A7.3 presents the estimates of Tobit models in which the dependent variable is either the amount sent by the trustor or the percentage returned by the trustee. They show that, even when controlling for covariates, the attitudes toward promoted subjects, either in the position of trustor or in the position of trustee, do not differ from the attitudes toward steady-lows.

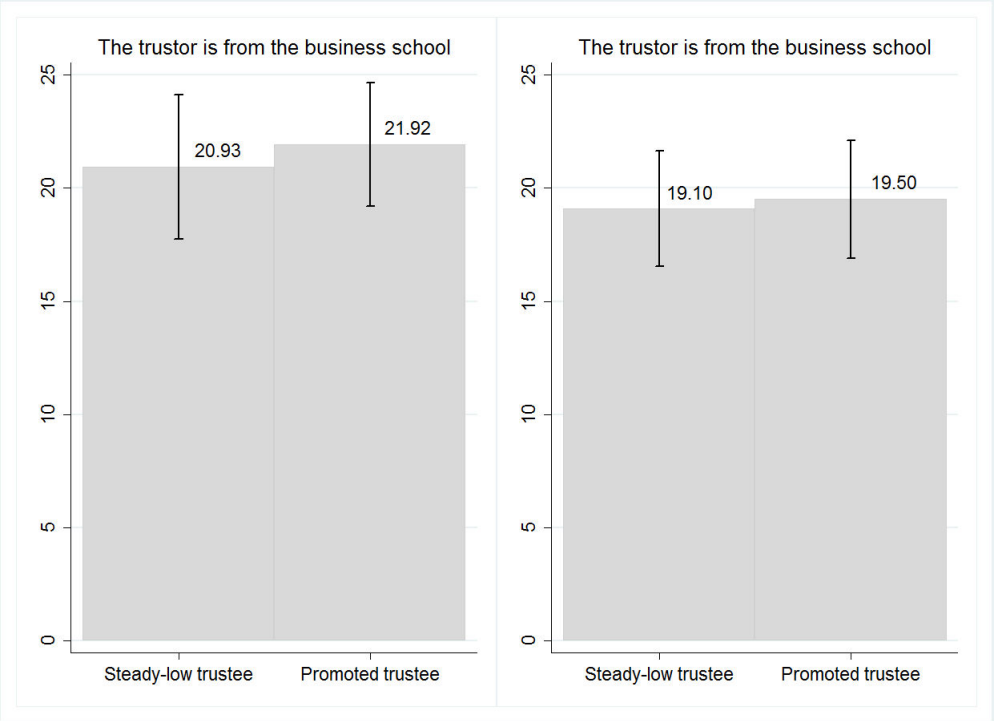


Figure 1.4 – Average percentage of the amount received returned to trustors from the business school, by statuts of the trustor and school of the trustee.

1.4.4 The role of inter-group comparison

In order to explore the role of inter-group comparison in the effect of upward mobility, we ran three sessions of an additional treatment, the one school condition, in which we recruited a total of 55 subjects from the business school only.[19] This treatment is very similar to the main treatment except that only two subjects in each group were promoted to be experts, instead of three in the main treatment. We did so in order to insure that promoted subjects are a minority, like in the main treatment. Because we have subjects from the business school only, we also replaced the incentives at the school level in the quiz about schools by a random payment. Every subjects in a session received an additional 5 ECU with probability one-half. Since all subjects come from the same school, we also removed the dictator game with subjects from the other school. This additional treatment is not directly comparable to the main treatment, which is why we analyze

[19]We thank an anonymous referee for suggesting this extension.

Table 1.4 – Comparison of the effect of promotion between the main condition and the one school condition

	(1) Trust Main. cond.		(2) Trust One school cond.	
Prom. trustor and Prom. trustee	-1.111	(1.126)	0.660	(1.910)
Prom. trustor and St. low trustee	-1.913**	(0.936)	0.437	(1.749)
St. low trustor and Prom. trustee	-0.473	(0.499)	0.137	(0.562)
Individual Char.	YES		YES	
Session F.E.	YES		YES	
N	174		220	
Pseudo R^2	0.111		0.126	
Log pseudo-likelihood	-370.152		-428.315	
F	4.88		5.11	
$p > F$	<0.001		<0.001	

Notes: Marginal effects are reported. Robust standard errors clustered at the individual level are in parentheses. ** $p < 0.05$, *** $p < 0.01$. (1) Data from the main condition. (2) Data from the one school condition.

Prom. denotes a promoted subject. St. low denotes a steady-low subject.

the data separately hereafter.

Table 1.4 reports marginal effect of Tobit regressions explaining trusting decision of subjects from the business school. The model estimated are identical to those estimated in Table 1.1, except that we cannot control for the transfer in the dictator game with a receiver from the other school, since all the subjects come from the business school in the one school condition. Model (1) uses the data from the main condition. Model (2) uses the data from the one school condition.

Result 1 is not replicated in the one school condition: when we remove group comparison, promotion does not reduce trust toward steady-lows. This suggests that inter-group comparison is necessary for promotion to impact trust.

The other two (null) results are replicated. Table A7.4 in A7 presents the results of regressions showing that promotion does not impact trustworthiness, and that promoted

individuals are not treated differently than steady-lows, nor in the main condition, neither in the condition with only one school.

1.5 Discussion and Conclusion

In this paper, our objective was to test experimentally the effect of upward social mobility on the trust and trustworthiness of promoted individuals and of others toward these individuals. We found limited evidence that trust and trustworthiness are affected by upward social mobility, defined in our experiment by the assignment of a position of expert to members of a natural group that is ex ante less likely to access this position compared to another natural group. The promoted individuals who reached the status of expert are as trustworthy and are not treated differently, in terms of trust or trustworthiness, than the other members of their natural group.

Our main result is that the promoted subjects tend to trust less both in-groups and out-groups. On the other hand, we find no effect of promotion on trustworthiness, irrespective of the identity of the trustor. These results are not consistent with our hypothesis according to which upward mobility should increase social distance between in-groups of different status and decrease social distance between out-groups of equal status. Indeed, this hypothesis implies (i) that promotion should reduce trust toward in-group but increase trust toward out-groups of high status, (ii) that trustworthiness should be impacted by promotion to the same extent as trust.

A possibility is that the high status might have a different meaning depending on whether it is held by a subject from the business school or a subject from the engineering school. In particular, promoted individuals may consider the status of expert as more rewarding for them than for subjects from the engineering school, since reaching the high status is not common for subjects from the business school, and since it was ex-ante harder for them due to their education. This possibility is supported by the data from

the one school condition, in which the specific experience of promoted subjects is not reinforced by the fact that their status is generally a characteristic of the out-group, and in which we find no effect of promotion on trust.

The merit of this explanation is that it also provides an explanation for the gap between the effect of promotion on trust and on trustworthiness. Trustors who hold a high status relative to the trustee have been shown to trust less, because they expect to suffer more from a potential betrayal (see e.g.Bohnet et al. (2008); Aimone and Houser (2012) on betrayal aversion and Hong and Bohnet (2007) for evidence linking relative status and betrayal aversion.). Our results are found controlling for beliefs and other individual characteristics, but not for one's betrayal aversion. As a consequence, betrayal aversion is a possible channel through which promotion may affect trust if promoted consider themselves of higher status than any non promoted individual. Betrayal aversion is relative to emotion regulation. Thus, emotional experience rather than social preferences or social distance could mediate the effect of promotion on interpersonal trust.

To sum up, the good news is that in our setting socially mobile individuals are still trusted by in-groups and out-groups who did not benefit from this mobility, and they remain as trustworthy as their in-groups. The bad news is that socially mobile individuals become less trusting. If our interpretation in terms of betrayal aversion is correct, this suggests the importance of accompanying social mobility by interventions to increase trusting behavior.

Of course, we must remain cautious before extrapolating our findings. Our design of social mobility is specific. First, being only based on performance in a quiz, the assignment of the status of expert is fair and transparent, whereas in real settings promotion processes are sometimes fuzzy or discretionary. This may potentially alter the relationships of the socially mobile individuals with others, as procedural fairness impacts social preferences (e.g., Bolton et al., 2005). Second, in our experiment, when subjects were

performing the quiz, they could not anticipate that their relative performance would be used to assign positions of experts and agents. The purpose was to limit selection into mobility on traits such as greed or preference for status, and to identify the causal effect of upward mobility on trust. In contrast, in real settings those who achieve social mobility are more likely to be attracted by competition, power and status compared to others. These preferences might affect the way upwardly mobile individuals treat others. For instance, Bartling et al. (2009) show that more competitive individuals are less egalitarian. These preferences may also affect how upwardly mobile individuals are perceived and thus, how much trust they inspire. Last, in contrast to our experiment which abstracts from income inequality, in real settings status inequality often comes with income inequality, which can also impact behavior toward mobile individuals (e.g. Lei and Vesely, 2010). The anecdotal evidence showing that socially mobile individuals face sometimes the risk of being rejected by their in-groups could result from these other considerations associated with upward mobility.

From a methodological point of view, we can also discuss about the procedures that were used to induce social mobility in the lab. Our subjects had both a natural affiliation and a lab-induced status. It could be that the school affiliation is so strong (which is supported by the in-group favoritism revealed by the dictator games) that the status induced was comparatively weak, and thus had a negligible impact on behavior. We doubt that this is the explanation of the observed behavior since we confirmed the importance of math for the students and we found that promotion reduced trust in non-promoted in-groups. However, we cannot exclude that our design may have induced status asymmetrically: promoted individuals may have perceived their status as specific, but others did not. This is consistent with the idea that individuals tend to identify more strongly with social categories that help them achieve a more positive image of themselves (see e.g. Hett et al., 2017).

Several interesting research directions could be explored. In particular, it would be important to study the impact of demotion on trusting behavior since social demotion is a rising concern in modern democracies. Moreover, promotion and demotion are often intertwined: one's promotion leads to someone else's demotion: how does this affect behavior toward promoted and demoted individuals? It would be also interesting to allow subjects to opt-in or opt-out of the promotion process to measure how selection into promotion affects trust toward promoted individuals. This opens an avenue for a new research program.

Bibliography

Acemoglu, D., G. Egorov, and K. Sonin (2017). Social mobility and stability of democracy: Re-evaluating de tocqueville. *The Quarterly Journal of Economics*.

Aimone, J. A. and D. Houser (2012). What you don't know won't hurt you: a laboratory analysis of betrayal aversion. *Experimental Economics 15*(4), 571–588.

Akerlof, G. A. and R. E. Kranton (2000). Economics and identity. *The Quarterly Journal of Economics 115*(3), 715–753.

Alesina, A. and E. L. Ferrara (2005). Preferences for redistribution in the land of opportunities. *Journal of Public Economics 89*(5-6).

Alesina, A., S. Stantcheva, and E. Teso (2017). Intergenerational mobility and support for redistribution. Working Paper 23027, National Bureau of Economic Research.

Arrow, K. J. (1972). Gifts and exchanges. *Philosophy and Public Affairs 1*(4), 343–362.

Austen-Smith, D. and R. G. Fryer (2005). An economic analysis of acting white. *The Quarterly Journal of Economics 120*(2), 551–583.

Bartling, B., E. Fehr, M. A. Maréchal, and D. Schunk (2009). Egalitarianism and competitiveness. *The American Economic Review 99*(2), 93–98.

Berg, J., J. Dickhaut, and K. McCabe (1995). Trust, reciprocity, and social history. *Games and Economic Behavior 10*(1), 122 – 142.

Bernhard, H., E. Fehr, and U. Fischbacher (2006). Group affiliation and altruistic norm enforcement. *American Economic Review 96*(2), 217–221.

Bock, O., I. Baetge, and A. Nicklisch (2014). Hroot: Hamburg registration and organization online tool. *European Economic Review 71*.

Bohnet, I., F. Greig, B. Herrmann, and R. Zeckhauser (2008). Betrayal Aversion: Evidence from Brazil, China, Oman, Switzerland, Turkey, and the United States. *The American Economic Review 98*(1), 294–310.

Bolton, G. E., J. Brandts, and A. Ockenfels (2005). Fair procedures: Evidence from games involving lotteries. *The Economic Journal 115*(506), 1054–1076.

Bowles, S. and H. Gintis (2004). Persistent parochialism: trust and exclusion in ethnic networks. *Journal of Economic Behavior & Organization 55*(1), 1–23.

Buchan, N. R., E. J. Johnson, and R. T. Croson (2006). Let's get personal: An international examination of the influence of communication, culture and social distance on other regarding preferences. *Journal of Economic Behavior & Organization 60*(3), 373–398.

Burks, S. V., J. P. Carpenter, L. Goette, and A. Rustichini (2009). Cognitive skills affect economic preferences, strategic behavior, and job attachment. *Proceedings of the National Academy of Sciences 106*(19), 7745–7750.

Charness, G., D. Masclet, and M. C. Villeval (2013). The dark side of competition for status. *Management Science 60*(1), 38–55.

Charness, G., L. Rigotti, and A. Rustichini (2007). Individual behavior and group membership. *American Economic Review 97*(4), 1340–1352.

Chen, R. and Y. Chen (2011). The potential of social identity for equilibrium selection. *American Economic Review 101*(6), 2562–2589.

Chen, Y. and S. X. Li (2009). Group identity and social preferences. *American Economic Review 99*(1), 431–457.

Chen, Y., S. X. Li, T. X. Liu, and M. Shih (2014). Which hat to wear? impact of natural identities on coordination and cooperation. *Games and Economic Behavior 84*, 58–86.

Corgnet, B., A. M. Espiìn, R. Hernán-González, P. Kujal, and S. Rassenti (2016). To trust, or not to trust: Cognitive reflection in trust games. *Journal of Behavioral and Experimental Economics 64*, 20–27.

Dohmen, T., A. Falk, D. Huffman, U. Sunde, J. Schupp, and G. G. Wagner (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association 9*(3), 522–550.

Dufwenberg, M. and A. Muren (2006). Generosity, anonymity, gender. *Journal of Economic Behavior & Organization 61*(1), 42–49.

Ellemers, N., F. Rink, B. Derks, and M. K. Ryan (2012). Women in high places: When and why promoting women into top positions can harm them individually or as a group (and how to prevent this). *Research in Organizational Behavior 32*, 163–187.

Etang, A., D. Fielding, and S. Knowles (2011). Does trust extend beyond the village? Experimental trust and social distance in Cameroon. *Experimental Economics 14*(1), 15–35.

Falk, A., A. Becker, T. Dohmen, B. Enke, D. Huffman, and U. Sunde (2018). Global evidence on economic preferences. *The Quarterly Journal of Economics*, qjy013.

Falk, A. and C. Zehnder (2013). A city-wide experiment on trust discrimination. *Journal of Public Economics 100*, 15–27.

Fershtman, C. and U. Gneezy (2001). Discrimination in a segmented society: An experimental approach. *The Quarterly Journal of Economics 116*(1), 351.

Fischbacher, U. (2007). Z-tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics 10*(2), 171–178.

Fryer, R. G. and P. Torelli (2010). An empirical analysis of acting white. *Journal of Public Economics 94*(5), 380–396.

Fukuyama, F. (1995). *Trust: The social virtues and the creation of prosperity*. Free Press Paperbacks.

Galeotti, F. and D. J. Zizzo (2014, feb). What happens if you single out? an experiment. *Social Choice and Welfare 43*(3), 703–729.

Goette, L., D. Huffman, and S. Meier (2006). The impact of group membership on cooperation and norm enforcement: Evidence using random assignment to real social groups. *American Economic Review 96*(2), 212–216.

Goldin, C. (2014). A pollution theory of discrimination: male and female differences in occupations and earnings. In *Human capital in history: The American record*, pp. 313–348. University of Chicago Press.

Heap, S. P. H. and D. J. Zizzo (2009). The value of groups. *American Economic Review 99*(1), 295–323.

Hett, F., M. Kröll, and M. Mechtel (2017). Choosing who you are: The structure and behavioral effects of revealed identification preferences. Technical report.

Hong, F., Y. Riyanto, and R. Zhang (2016). Multidimensional group identity and redistributive allocation: An experimental study. Economic Growth Centre Working Paper Series 1701, Nanyang Technological University, School of Social Sciences, Economic Growth Centre.

Hong, K. and I. Bohnet (2007). Status and distrust: The relevance of inequality and betrayal aversion. *Journal of Economic Psychology 28*(2), 197–213.

Johnson, N. D. and A. A. Mislin (2011). Trust games: A meta-analysis. *Journal of Economic Psychology 32*(5), 865–889.

Kazdin, A. E. (2000). *Encyclopedia of Psychology*. Oxford University Press, New York.

Klor, E. F. and M. Shayo (2010). Social identity and preferences over redistribution. *Journal of Public Economics 94*(3), 269–278.

Kranton, R., M. Pease, S. Sanders, and S. Huettel (2016). Group bias, identity, and social preferences. Technical report, Working Paper.

Kuhn, P. (2017). *Personnel economics*. Oxford University Press.

LaPorta, R., F. L. de Silanes, A. Shleifer, and R. W. Vishny (1997). Trust in large organizations. *American Economic Review Papers and Proceedings 87*(2), 333–338.

Lazear, E. P. and M. Gibbs (2014). *Personnel economics in practice*. John Wiley & Sons. 416p.

Lei, V. and F. Vesely (2010). In-group versus out-group trust: The impact of income inequality. *Southern Economic Journal 76*(4), 1049–1063.

Menger, P.-M. and C. Marchika (2014). La technologie de sélection des étudiants dans les grandes écoles de commerce françaises. *Revue Française d'Economie 0*(2), 5–42.

Oakley, J. G. (2000). Gender-based barriers to senior management positions: Understanding the scarcity of female CEOs. *Journal of Business Ethics 27*(4), 321–334.

Piketty, T. (1995). Social mobility and redistributive politics. *The Quarterly Journal of Economics 110*(3), 551–584.

Prendergast, C. (1999). The provision of incentives in firms. *Journal of Economic Literature 37*(1), 7–63.

Sen, A. (1980). *Equality of What?* Cambridge: Cambridge University Press.

Stanca, L., L. Bruni, and L. Corazzini (2009). Testing theories of reciprocity: Do motivations matter? *Journal of Economic Behavior & Organization 71*(2), 233 – 245.

Tajfel, H. and J. C. Turner (1979). An integrative theory of intergroup conflict. *The Social Psychology of Intergroup Relations 33*(47), 74.

Tsutsui, K. and D. J. Zizzo (2013). Group status, minorities and trust. *Experimental Economics 17*(2), 215–244.

A1 Timeline of the experiment

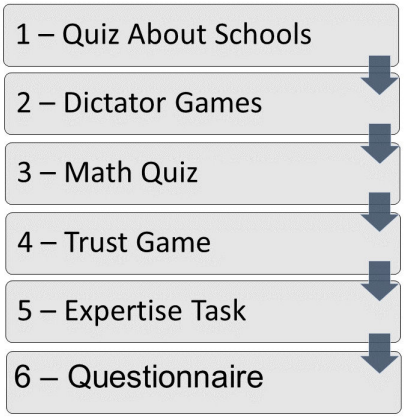


Figure A1.1 – Timeline of the experiment

A2 Instructions *(Translated from French)*

We thank you for participating in this experiment on decision-making. Please switch off you cellphone and put it away. You are not allowed to communicate with the other participants, unless otherwise instructed by an experimenter.

During the session, if you have any question you can press the red button on the side of your cubicle. An experimenter will come and answer to your questions in private. During the session, you will have to make several decisions anonymously. These decisions can earn you money. Your earnings will be expressed in Experimental Currency Units (ECU) and converted into Euros at the following rate:

5 ECU = €1

You will be paid in private, in a separate room and in cash. Other participants will not be informed of your earnings.

The session consists of several parts. At the end of each part, you will receive the instructions for the next part. In this experiment, participants are students from the Ecole Centrale de Lyon and from the Ecole de Management de Lyon. We call participants from the Ecole Centrale "Centraliens" and participants from the Ecole de Management "Emiens". Please make sure that the logo displayed on your computer screen corresponds to your school.

Part 1

In the first part, you have to answer individually to a quiz about your school. 6 multiple choice questions will be displayed on your screen. For each question, you have to choose an answer and validate by pressing the OK button. Once you have pressed the ok button, your answer is recorded and you proceed to the next question. You have 4 minutes to answer to the 6 questions.

In order to get help to answer to the questions, you can use a chatbox, displayed on the right part of the screen. You can communicate only with the participants from the same school as you and exclusively through the chatbox. Communication is anonymous. You can send any message, provided that these messages do not identify you and are not offensive.

In this part, you will earn a fixed payoff of 5 ECU and a variable payoff that depends on your answers. Each correct answer pays you 2.5 ECU. In addition, each participant from the school whose participants in the session gave the highest number of correct answers earns an extra 5 ECU.

You will be informed of your number of correct answers, of the school which participants gave the highest number of correct answers and of your payoff in this part at the end of the session.

***** Please read these instructions again. If you have any question, press the call button on the side of your cubicle *****

Part 2

This part involves person A and person B.

Person A receives an endowment of 10 ECU. He has to decide which amount, between 0 and 10 ECU inclusive, he is willing to transfer to person B. He keeps for himself the amount he did not transfer.
Person B does not receive any endowment. He earns the amount that person A transfers to him. He has no decision to make.
Each participant makes two decisions successively as a person A : in one decision, person B is a student from Ecole Centrale de Lyon; in the other decision, person B is a student from Ecole de Management de Lyon.

At the end of the session, the computer program will randomly match you to two other participants. For one of these matches, you will be paid for your decision as person A; for the other match, you will be payed in the role of person B. The program selects randomly the match for which your decision as person A determines your payoff.

1. As a person A, it is your decision that determines your payoff and the payoff of the person B.
2. As a person B, it is the decision of the person A you are matched with that determines your payoff and his payoff.

You will be informed of your payoff in this part at the end of the session.

***** Please read these instructions again. If you have any question, press the call button. *****

Part 3

In this part, the computer program forms randomly groups of five participants. In each group, two participants are from Ecole Centrale de Lyon and three participants are from Ecole de Management de Lyon.

You have to answer individually to multiple choice mathematical questions. You have to select an answer and validate it by pressing the OK button. Validation is definitive. You can use the paper sheets and pen that have been provided to you. You are not allowed to use your cellphone to help you solve the questions, otherwise you expose yourself to exclusion from the session and from the payoffs. Every

participants in the session receive the same questions in the same order.

You have 15 minutes to answer to the questions.

For your participation in this part, you earn a fixed payoff of 5 ECU and a variable payoff. The computer program will randomly select one participant in each group of five participants. The number of correct answers of this participant will determine the variable payoff of each member of his group. Each correct answer of this participant increases the payoff of every member of his group, including himself, by 1 ECU.

You will be informed of your number of correct answers and of your payoff in this part at the end of the session. In addition, your relative performance in this part will condition your role in the next part.

***** Please read these instructions again. If you have any question, press the call button. *****

Part 4

The previous part was used by the computer program to identify two types of participants who will have different roles in what follows.

1. Experts are participants who gave the highest number of correct answers in the third part within their group of five participants. In each group of five, there are three experts.
2. Agents are the participants who are not experts. In each group of five persons, there are two agents.

You are informed whether you are an agent or an expert at the beginning of the fourth part. This part consists in two stages.

Stage 1

In this stage, the composition of the groups of five is the same as in the third part. This stage consists in eight successive games. In each game, you are paired with a different member of your group. You are informed of the school and role (expert or agent) of the other member of the pair.

In each pair, a participant is a participant A and the other one is a participant B. The sequence in each game is the following :

1. Participant A makes a decision.
2. When making his decision, participant B does not know the decision made by participant A. Participant B has to make a decision for each potential decision made by participant A.

In the first four games, you have the role of person A and you interact with the four other members of the group in the role of person B.

In the last four games, you have the role of person B and you interact with the four other members of the group who will have the role of person A.

At the end of the session, the computer program will randomly select one of your decisions in the role of person A and one of your decision in the role of person B. Your payoff in each of these two games will be added-up to determine your payoff in this part.

Description of each game:

1. Participant A and participant B receive an endowment of 10 ECU each.
2. Participant A chooses the amount he is willing to send to participant B. Participant A can send from 0 to 10 ECU, inclusive.
3. Each ECU sent to participant B is multiplied by 3 by the computer program. For example, if participant A sends 2 ECU, participant B receives 2x3=6 ECU ; if he sends 4 ECU, participant B receives 4x3=12 ECU.
4. Then, participant B chooses the amount he is willing to return to participant A. This amount is between 0 and three times the amount sent by participant A, inclusive.

When choosing the amount to return to participant A, participant B does not know the amount sent by participant A. Participant B has to choose the amount he is willing to send back to participant A for each amount participant A potentially sent to him. For each amount potentially sent by participant A, participant B can return any amount between 0 and 3 times this amount (because he received this amount multiplied by 3). For example, if participant A sent 2 ECU, participant B can send back any amount, between 0 and 6 ECU, inclusive. If participant A sent 5 ECU, he can send back any amount, between 0 and 15 ECU, inclusive.

When choosing which amount to send to participant B, participant A has to indicate which proportion of the amount received by participant B he expects to receive in return. A guess equal to the actual amount more or less 5 percents pays an extra 5 ECU.

Determination of payoffs For each game selected for payment, the computer program takes into account the decision of Participant A. Then, the program selects among participant B's return decisions the one that corresponds to the amount actually sent by participant A.

For each game selected for payment in this stage, participant A's payoff is computed as follows:

Payoff of participant A = 10 - amount sent to participant B + amount sent back by participant B

Participant B's payoff is computed as follows:

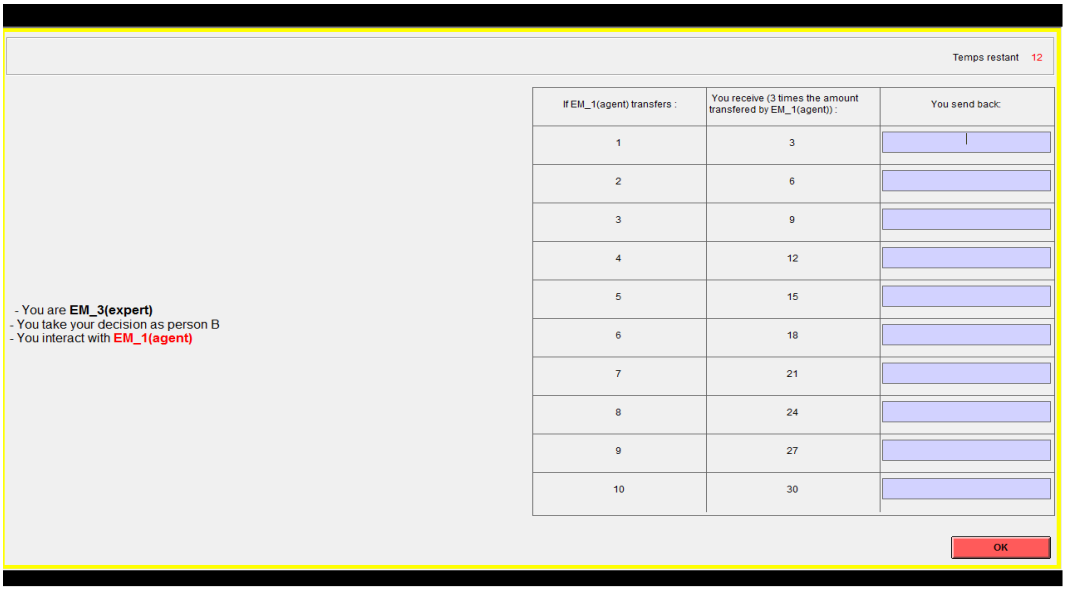
Payoff of participant B =10 + 3*amount sent by A - amount sent back to A

The following figure represents the screenshot for the decision of a participant A. On this screen, you have to indicate the amount you are willing to send to participant B and the proportion of the amount received you guess you will receive in return.

Figure A2.1 – Screenshot of participant A's decision.

The following figure represents the screenshot for the decisions of a participant B. The first column indicates each amount potentially sent by A. The second column displays each corresponding tripled amount you can potentially receive. In the third column you have to enter on each line the amount you decide to send back to participant A, between 0 and the tripled amount indicated in the second column.

Figure A2.2 – Screenshot of participant B's decisions.



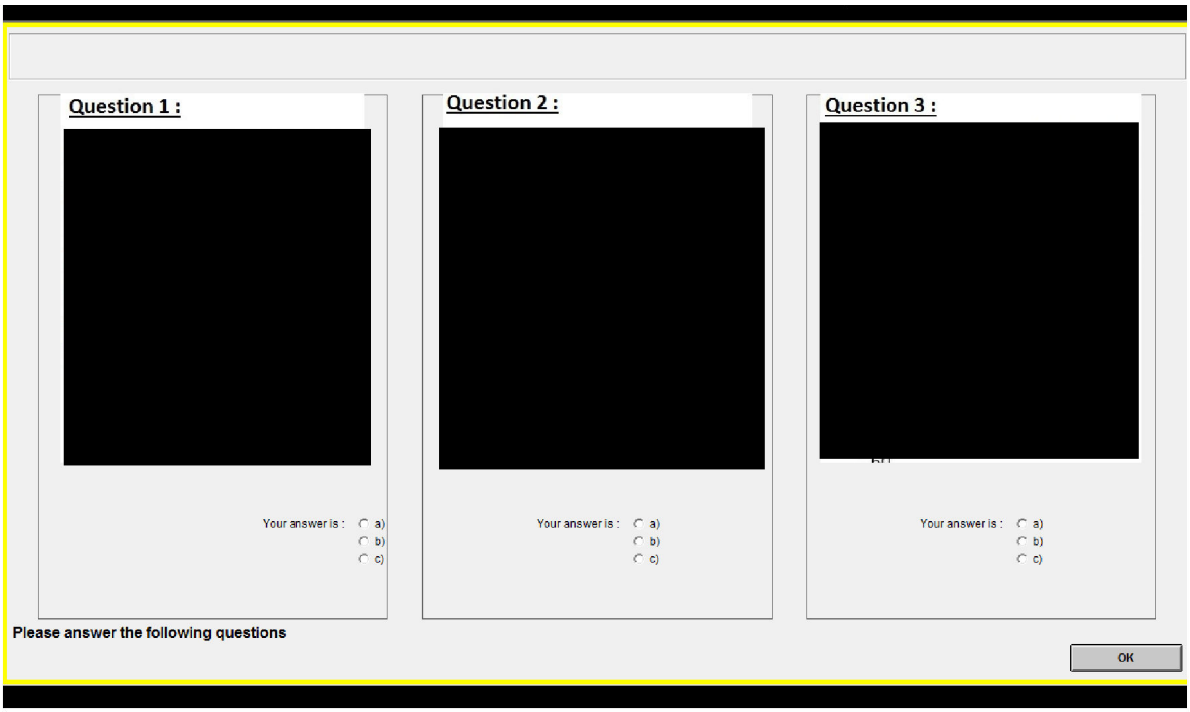
Stage 2

In the second stage, decisions are no longer made within the five person group, but at the session level.

Agents and experts have different roles. Agents have to answer to three multiple choice maths questions. Each agent receives the three same questions. Experts are informed of the distribution of the agents' answers; then, they have to decide which question will be used to determine everybody's payoff in this part.

The following screenshot represents the agents' decision screen.

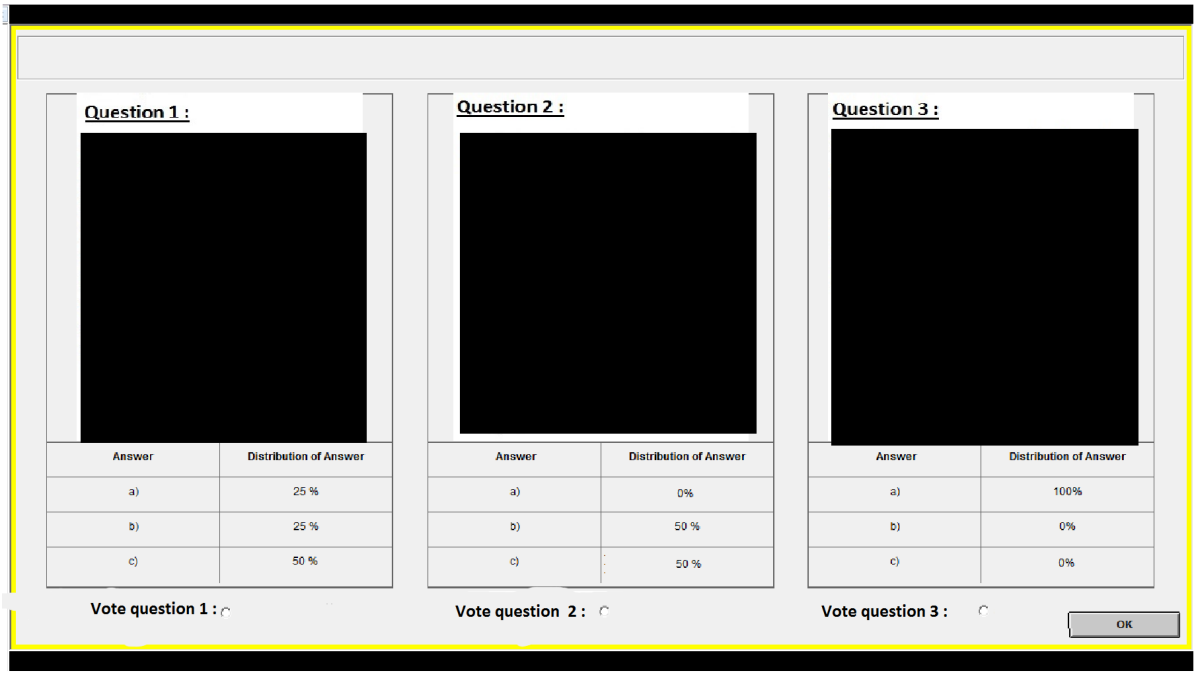
Figure A2.3 – Screenshot of the Agents' decisions



Once all the agents have submitted their answers, the experts can see the questions and the distribution of the answers for each question. Experts have to chose which question will be used to determine the payoff of each participant in the session for this stage, regardless of their role of agent or expert. To do so, they vote for one of the questions. The number of correct answers to the question that has received the highest number of votes from the experts in the session determines each participant's payoff in this stage.

The following screenshot represents the experts' decision screen.

Figure A2.4 – Screenshot of the Experts' decision



The payoff in ECU of each participant in this stage is equal to the percentage of correct answers to the selected question, multiplied by 10. For instance, if the experts choose question 1 and that 50% of the agents gave the correct answer to this question, each participant in the session earns 5 ECU (50%*10 =5).

***

At the beginning of this part, you will be informed on whether you are an expert or an agent. You will receive a unique identifier in your group of five persons in the form: School_i(expert) or School_i(agent). Then, you will have to fill out a check questionnaire that will be displayed on your screen, then stage 1 will start.

***

At the end of this part, you will be informed of your performance and of your payoffs in the different parts of the experiment. At the end of the session, several questionnaires will be displayed on your screen. We remind you that your answers are anonymous. Once you have filled these questionnaires, please remain seated and silent. When you are called to the payment room, bring with you only your computer tag and your payment receipt. Please, leave the instructions, pen and papers on your desk.

**** Please read these instructions again and if you have questions, press the call button. *****

A3 Quiz relative to schools used in Part 1 *(Translated from French)*

Business school

Question 1 How many students are there at EM Lyon?

1. Around 1000
2. Around 2000
3. Around 3000
4. Around 4000

Question 2 According to the ranking by "L'etudiant",[20], what is the rank of EM Lyon among the French business schools?

1. Third rank
2. Fifth rank
3. Tenth rank
4. Beyond tenth rank

Question 3 Among the following famous people, which one is an alumni from EM Lyon?

1. Nagui
2. Stéphane Bern
3. Christophe Dechavanne
4. Julien Courbet

Question 4 When was EM Lyon founded?

1. 1753
2. 1872
3. 1917
4. 1932

Question 5 What is the average grade at Baccalaureat[21] of the students at EM Lyon?

1. 15
2. 15.5
3. 16.5
4. 17.5

Question 6 What is the average yearly salary of the students from EM Lyon in their first position after graduation?

1. Less than €30 000
2. Between €30 and €35 000
3. Between €35 and €40 000
4. More than €40 000

[20] L'Etudiant is a magazine about higher education that is widely read by students.
[21] In France, Baccalaureat is the exam passed at the end of the high school.

Engineering school

Question 1 How many students are there at ECL?

1. Less than 1000
2. About 1500
3. About 2000
4. About 2500

Question 2 According to the ranking by "L'etudiant", what is the rank of ECL among the French engineering schools?

1. Less than tenth
2. Between tenth and fifteenth
3. Between fifteenth and twentieth
4. Beyond twentieth

Question 3 Among the following famous people, which one is an alumni from ECL?

1. Jean Mermoz
2. Paul-Emile Victor
3. Jacques-Yves Cousteau
4. Nicolas Hulot

Question 4 When was ECL founded?

1. 1753
2. 1857
3. 1912
4. 1934

Question 5 What is the average grade at Baccalaureat of the students at ECL?

1. 15
2. 16
3. 17
4. 18

Question 6 What is the average yearly salary of the students from ECL in their first position after graduation?

1. Less than 27 000€
2. Between 27 and 30 000€
3. Between 30 and 33 000€
4. More than 33 000€

Figure A3.1 shows a screenshot of the quiz about schools.

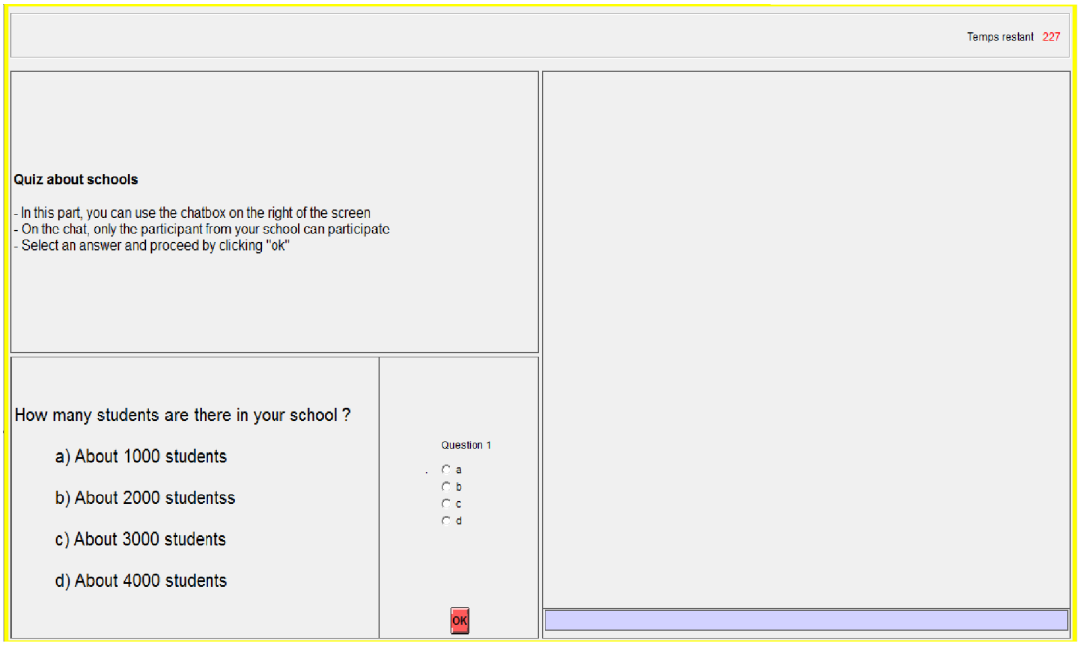


Figure A3.1 – Example of a screenshot of the quiz about schools.

A4 Math quiz used to assign positions of Agents and Experts in Part 3 *(Translated from French)*

1. a and b are two even integers. Which of the following is also an even integer?
   (a) $ab + 2$
   (b) $a(b - 1)$
   (c) $a(a + 5)$
   (d) $3a + 4b$
   (e) $(a + 3)(b - 1)$
2. $(x + 2)^2 = -4 + 10x$. What can be a value of x?
   (a) 2
   (b) 1
   (c) 0
   (d) -1
   (e) -2
3. Approximately, what percentage of the forest across the world is in Finland, knowing that Finland has 53.42 millions hectares of forest for a total of 8.076 billions hectares of forest worldwide?
   (a) 0.0066%
   (b) 0.066%
   (c) 0.66%
   (d) 6.6%
   (e) 66%
4. Figure A4.1 is a square. This square has sides that are 4 units long. What is the best approximation of the area of the circle?
   (a) π
   (b) 4
   (c) 8
   (d) 13
   (e) 16
5. An individual invests his money in stocks in the financial market. During the first year, the value of his stocks increases by 50%. During the second year, the value of his stocks decreases by 30%. What is the total variation of the value of his stocks in the period?
   (a) -5%
   (b) 5%
   (c) 15%
   (d) 20%
   (e) 80%
6. Assume that A, B, C are three statements such that C is true if exactly one of A or B is true. If C is false, then which of the following statement is necessarily true?
   (a) If A is true, then B is false
   (b) If A is false, then B is false
   (c) If A is false, then B is true
   (d) A and B are both true
   (e) A and B are both false
7. $(1 + i)^{10} =$?
   (a) 1
   (b) i
   (c) 32

(d) $32i$

(e) $32(i+1)$

8. f is a real value function continuously differentiable, defined on the open interval $(-1, 4)$ such that $f(3) = 5$ and $f'(x) \leq -1$ for all x. What is the greatest possible value of $f(0)$?

(a) 3

(b) 4

(c) 5

(d) 8

(e) 11

9. A drawer contains 2 pairs of blue socks, 4 pairs of red socks, 2 pairs of yellow socks. If we draw randomly two pairs of socks from this drawer, what is the probability that those two pairs are of the same color?

(a) $\frac{2}{7}$

(b) $\frac{2}{5}$

(c) $\frac{3}{7}$

(d) $\frac{1}{2}$

(e) $\frac{3}{5}$

10. If $F(x) = \int_e^x ln(t)dt$ for all x, then $F'(x) =$?

(a) x

(b) $\frac{1}{x}$

(c) $ln(x)$

(d) $xln(x)$

(e) $xln(x) - 1$

11. $F(1)$ and $F(n) = F(n-1) + \frac{1}{2}$ for all integer $n > 1$, then $F(101) =$?

(a) 49

(b) 50

(c) 51

(d) 52

(e) 53

12. $lim_{x\to 0}\frac{cos(3x)-1}{x^2} =$?

(a) $\frac{9}{2}$

(b) $\frac{3}{2}$

(c) $\frac{-2}{3}$

(d) $\frac{-3}{2}$

(e) $\frac{-9}{2}$

13. Assume that f is differentiable, with $lim_{x\to\infty}f(x)$ and $lim_{x\to\infty}f'(x)$ both existing and finite. Which of the following statement MUST be true?

(a) $lim_{x\to\infty}f'(x) = 0$

(b) $lim_{x\to\infty}f''(x) = 0$

(c) $lim_{x\to\infty}f(x) = lim_{x\to\infty}f'(x)$

(d) f is constant

(e) f' is constant

14. What is the $19th$ derivative of $\frac{x-1}{e^x}$?

(a) $(18 - x)e^{-x}$

(b) $(19 - x)e^{-x}$

(c) $(20 - x)e^{-x}$

(d) $(x - 19)e^{-x}$

(e) $(x - 20)e^{-x}$

15. How many positive numbers satisfy $cos(97x) = x$?

(a) 1

(b) 15

(c) 31

(d) 49

(e) 96

16. $\Sigma_{k=1}^{\infty} \frac{k^2}{k!} = ?$

(a) e

(b) $2e$

(c) $(e+1)(e-1)$

(d) e^2

(e) ∞

17. The first derivative of $\phi(t) = f(t^2, 2t)$ is:

(a) $(2+2t)\frac{\partial f}{\partial x}(t^2, t2) + (2+2t)\frac{\partial f}{\partial y}(t^2, 2t)$

(b) $\frac{\partial f}{\partial x}(t^2, 2t) + \frac{\partial f}{\partial y}(t^2, 2t)$

(c) $2t\frac{\partial f}{\partial x}(t^2, 2t) + 2\frac{\partial f}{\partial y}(t^2, 2t)$

(d) $2\frac{\partial f}{\partial x}(t^2, 2t) + 2t\frac{\partial f}{\partial y}(t^2, 2t)$

(e) None of the previous is correct

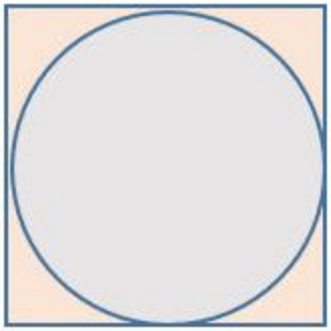


Figure A4.1 – Square and circle

Figure A4.2 shows a screenshot of the math quiz.

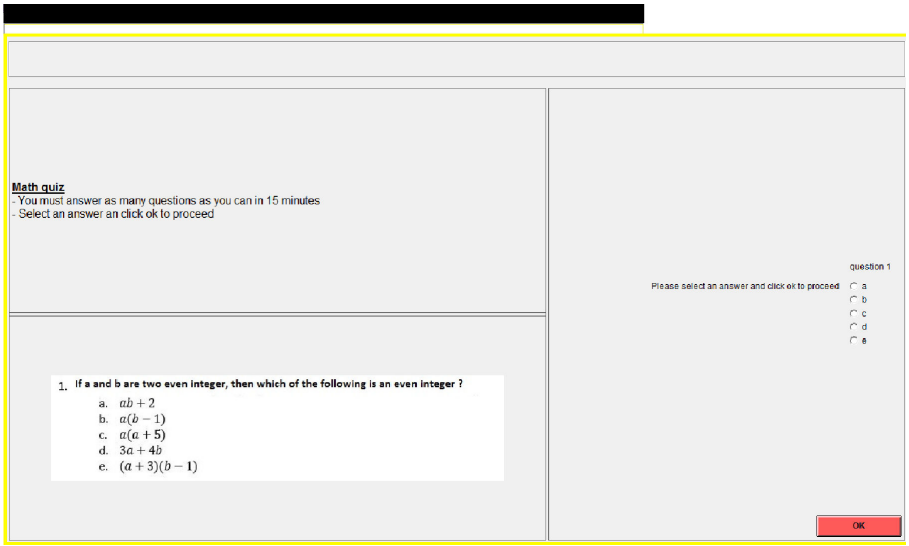


Figure A4.2 – Example of a screenshot of the math quiz

A5 Robustness test 1

In the main condition, we informed the subjects about the content of the expertise task before the trust game. A concern is that it may confound our results: the common payoff in the expertise task could generate solidarity or reciprocity between experts and agents. In order to asses the severity of this concern, we ran three sessions of a robustness test in which information about the nature the expertise task given before the trust game is minimal. A total of 45 subjects participated in this condition. In this robustness test, we only inform subjects that there will be a task in which experts and agents will have different roles after the trust game and that payoffs will be equal across status, but we do not inform them of the precise content of the task or how payoffs will be determined. In this treatment, solidarity or reciprocity concerns cannot stem from anticipation about the content of the expertise task.

In order to formally test whether our main results hold across the main and the robustness conditions, we ran models similar to model (2) and (3) in Table 1.1, except that we have interacted the variables of interest indicating identity (Prom. trustee and st.low trustor, Prom. trustee and st. high trustee etc.) with a dummy variable indicating the robustness condition (robust. cond.). We performed Chow tests comparing the parameter supporting our results across conditions. Table A5.1 reports the outcome of the regression and of the tests. Model (1) and (2) explain trust decisions of subjects from the business school. In Model (1), both the trustor and the trustee are from the business school. In model (2), the trustee is from the engineering school. Model (3) and (4) explain trustworthiness of subjects from the business school. In Model (3), both the trustor and the trustee are from the business school. In model (4), the trustor is from the engineering school. We report the $p - value$ of Chow tests comparing the parameter of interest across conditions.

Table A5.1 – Trust and trustworthiness of subjects from the business school across the main and the robustness condition

	(1) Trust		(2) Trust		(3) Trustworthiness		(4) Trustworthiness	
Prom. sender and St. low receiver[a]	-2.383**	(1.126)			3.827	(9.461)		
Prom. sender and St. low receiver (rob. cond.)	-1.062	(1.372)			6.636	(15.98)		
Prom. sender and St. high receiver			-2.079*	(1.068)			9.763	(7.890)
Prom. sender and St. high receiver (rob. cond.)			-2.092	(1.476)			3.279	(14.63)
$p-value$ diff.	0.474		0.994		0.879		0.695	
Individual Char.	YES		YES		YES		YES	
Session F.E.	YES		YES		YES		YES	
N	174		174		174		174	

Notes: Marginal effects are reported. Robust standard errors clustered at the individual level are in parentheses. ** $p < 0.05$, *** $p < 0.01$. (1) Trustee from the business school only, (2) Trustee from the engineering school only (3) Trustor from the business school only, (4) Trustor from the engineering school only. (a) Sender refers to the trustor when the explained decision is trust, and to the trustee when the explained decision is trustworthiness. $p-value$ diff. corresponds to the $p-value$ of Chow tests comparing parameters on two successive lines.

In each case, we find that the parameters supporting our results are not statistically different across conditions. As a consequence, we conclude that informing subjects of the content of the expertise task before the trust game does not drive our results. We pool the data from the robustness condition together with the data of the main condition.

A6 Robustness test 2

In this part, we further check that differences in math abilities do not explain Result 1. In Table 1.1, we controlled for math abilities in order to account for the effect of math abilities on trust. In this robustness check, rather than controlling for the score in the math quiz in the regressions, we leverage a feature of our design: status is awarded in groups of five to which subjects are randomly assigned. As a consequence, there is a large overlap between the performances of experts and agents from the business school as shown in Figure A6.1. The minimal score for a subject from the business school to be an expert was 4. There are 69 (79%) subjects who scored 4 or more, among which 43 experts and 26 agents. Each of this subjects could have been awarded the status of expert had they been randomly assigned to a different group.

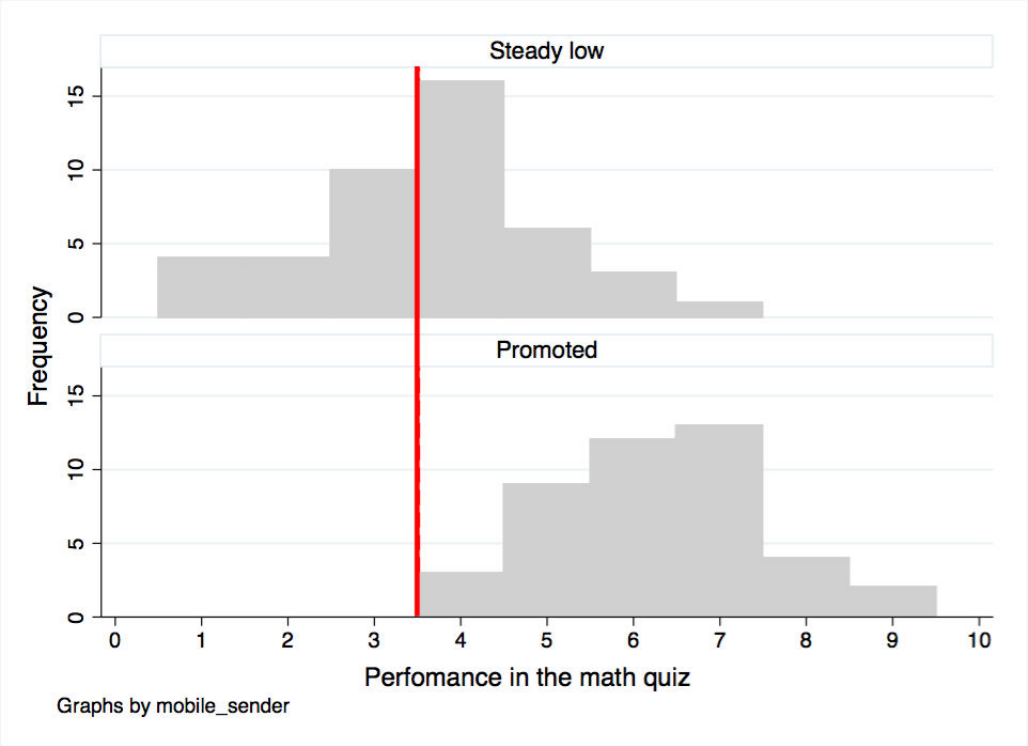


Figure A6.1 – Distribution of performance of the subjects from the business school in the math quiz, by status

In Table A6.1, we estimate a model similar to the model (1) in Table 1.1 excepted that rather than controlling for math abilities thanks to the score in the math quiz, we estimate the model on the subset of subjects who scored 4 or more, i.e. the subset of subjects who are "good enough" to be experts in at least one group. Some of this subjects are experts and other are agents.

The results from model (1) in Table A6.1 are qualitatively in line with those presented in model (1) in Table 1.1. Significance level are lower in this models because we drop some observations from the estimation. This suggests that our results are robust to various ways of controlling for math abilities.

A7 Additional Regressions

Table A6.1 – Trust by subjects from the business school who scored 3 or more in the math quiz..

	Tobit models					
	(1)		(2)		(3)	
	Trust		Trust		Trust	
Beliefs	0.0686****	(0.0171)	0.0814****	(0.0197)	0.0617***	(0.0195)
Prom. trustor and Prom. trustee	-0.671	(0.948)	-0.514	(0.859)	-	-
Prom. trustor and St. low trustee	-1.431*	(0.769)	-1.320*	(0.731)	-	-
Prom. trustor and St. high trustee	-1.956**	(0.834)	-	-	-1.531	(1.382)
Prom. trustor and Dem. trustee	-1.193	(0.934)	-	-	-0.850	(1.558)
St. low trustor and Prom. trustee	0.0262	(0.721)	0.0911	(0.656)	-	-
St. low trustor and St. high. trustee	-0.547	(0.744)	-	-	-0.166	(1.071)
St. low trustor and Dem. trustee	-0.414	(1.108)	-	-	-	-
Transfer in the DG, same school	0.166	(0.171)	0.189	(0.156)	0.114	(0.197)
Transfer in the DG, other school	0.445**	(0.225)	0.305	(0.231)	0.564**	(0.232)
Female	-0.205	(0.693)	-0.288	(0.702)	-0.126	(0.782)
Risk attitude[a]	0.0559	(0.143)	0.234	(0.158)	-0.142	(0.181)
Perception Math[b]	-1.076	(0.666)	-1.306*	(0.692)	-0.922	(0.703)
Session F.E.	YES		YES		YES	
N	276		138		138	
Pseudo R^2	0.125		0.160		0.116	
Log pseudolikelihood	-559.387		-273.381		-276.487	
F	6.04		7.84		4.98	
$p > F$	< 0.001		< 0.001		< 0.001	

Notes: Marginal effects are reported. Robust standard errors clustered at the individual level are in parentheses. ** $p < 0.05$, *** $p < 0.01$. (1) Trustees from the business school and from the engineering school pooled, (2) Trustees from the business school only, (3) Trustees from the engineering school only.

The reference category in models (1) and (2) corresponds to steady-low trustors matched with a steady-low trustee; in model (3), it is a steady-low trustor matched with an steady-high trustee. Prom. for promoted (business school), St. low for steady-low (business school), steady-high for steady-high (engineering school), and Dem. for demoted (engineering school). (a) For risk attitudes, a higher number indicates less risk aversion. (b) Dummy indicating that the trustor chose "strongly disagree" or "disagree" about the statement *"It is important to have good mathematical skills"*.

Table A7.1 – Beliefs of the Trustors from the Business School about the Percentage Returned by the Trustee

	Tobit models					
	(1) Beliefs		(2) Beliefs		(3) Beliefs	
Prom. trustor and Prom. trustee	-4.234	(10.12)	-4.151	(10.48)	-	-
Prom. trustor and St. low trustee	-3.917	(10.24)	-1.526	(9.744)	-	-
Prom. trustor and St. high trustee	5.180	(10.85)	-	-	3.007	(11.02)
Prom. trustor and Dem. trustee	-6.231	(10.12)	-	-	-5.900	(11.13)
St. low trustor and Prom. trustee	4.649	(4.151)	3.473	(3.843)	-	-
St. low trustor and St. high. trustee	-0.0750	(4.477)	-	-	-	-
St. low trustor and Dem. trustee	8.937	(10.92)	-	-	11.26	(9.214)
Transfer in the DG, same school	1.982	(1.340)	2.282	(1.411)	1.687	(1.541)
Transfer in the DG, other school	4.246**	(1.661)	4.566**	(1.790)	3.852**	(1.789)
Female	-13.17*	(7.285)	-12.10	(7.656)	-14.19*	(8.332)
Risk attitude[a]	1.237	(1.394)	1.815	(1.444)	0.541	(1.739)
Math quiz performance	-1.074	(3.099)	-2.102	(3.505)	-0.139	(3.304)
Perception Math[b]	-6.139	(11.55)	-4.476	(10.57)	-7.725	(13.64)
Session F.E.	YES		YES		YES	YES
N	348		174		174	
Pseudo R^2	0.0359		0.048		0.033	
Log pseudolikelihood	-1245.716		-607.442		-631.964	
F	3.48		3.59		2.47	
$p > F$	< 0.001		< 0.001		0.001	

Notes: Marginal effects are reported. Robust standard errors clustered at the individual level are in parentheses. ** $p < 0.05$. (1) Trustees from the business school and from the engineering school pooled, (2) Trustees from the business school only, (3) Trustees from the engineering school only.

The reference category in models (1) and (2) corresponds to steady-low trustors matched with a steady-low trustee; in model (3), it is a steady-low trustor matched with an steady-high trustee. Prom. for promoted (business school), St. low for steady-low (business school), steady-high for steady-high (engineering school), and Dem. for demoted (engineering school). (a) For risk attitudes, a higher number indicates less risk aversion. (b) Dummy indicating that the trustor chose "strongly disagree" or "disagree" about the statement *"It is important to have good mathematical skills"*.

Table A7.2 – Trusworthiness of Trustees from the Business School

	Tobit models					
	(1) Trustworthiness		(2) Trustworthiness		(3) Trustworthiness	
Trust	3.110****	(0.733)	3.106****	(0.900)	3.331****	(0.767)
Prom. trustor and Prom. trustee	-3.833	(8.347)	-3.984	(8.663)		
Prom. trustor and St. low trustee	0.574	(7.990)	-0.514	(8.490)		
Prom. trustor and St. high trustee	-2.484	(7.940)			3.975	(7.311)
Prom. trustor and Dem. trustee	-9.998	(9.124)			-4.502	(8.893)
St. low trustor and Prom. trustee	-1.753	(3.495)	-1.755	(3.436)		
St. low trustor and St. high. trustee	-4.910	(3.127)				
St. low trustor and Dem. trustee	-5.567	(7.091)			-0.314	(5.631)
Transfer in the DG, same school	1.760*	(0.974)	1.590	(1.092)	1.871**	(0.946)
Transfer in the DG, other school	3.510**	(1.456)	3.676**	(1.654)	3.278**	(1.410)
Female	-5.855	(5.032)	-5.480	(5.424)	-6.255	(5.141)
Risk attitude[a]	0.361	(1.256)	-0.0383	(1.403)	0.885	(1.223)
Math quiz performance	-0.606	(1.782)	-0.111	(1.993)	-1.156	(1.816)
Perception Math[b]	1.146	(7.563)	2.626	(8.403)	-0.360	(7.273)
Session F.E.	YES		YES		YES	
N	348		174		174	
Pseudo R^2	0.069		0.064		0.073	
Log pseudolikelihood	-1076.697		-552.113		-522.120	
F	4.28		3.74		4.61	
$p > F$	0.000		0.000		0.000	

Notes: Marginal effects are reported. Robust standard errors clustered at the individual level are in parentheses. ** $p < 0.05$, *** $p < 0.01$. (1) Trustors from the business school and from the engineering school pooled, (2) Trustors from the business school only, (3) Trustors from the engineering school only.

Notes: Marginal effects are reported. Robust standard errors clustered at the individual level are in parentheses. ** $p < 0.05$. (1) Trustees from the business school and from the engineering school pooled, (2) Trustees from the business school only, (3) Trustees from the engineering school only.

The reference category in models (1) and (2) corresponds to a steady-low trustee matched with a steady-low trustor; in model (3), it is a steady-low trustee matched with an steady-high trustor. Prom. for promoted (business school), St. low for steady-low (business school), steady-high for steady-high (engineering school), and Dem. for demoted (engineering school). (a) For risk attitudes, a higher number indicates less risk aversion. (b) Dummy indicating that the trustor chose "strongly disagree" or "disagree" about the statement *"It is important to have good mathematical skills"*.

Table A7.3 – Trust and Trustworthiness Toward Promoted Counterparts

	Tobit models							
	(1) Trust		(2) Trust		(3) Trustworthiness		(4) Trustworthiness	
Trust					3.106****	(0.900)	2.523***	(0.966)
Beliefs	0.0521***	(0.0175)	0.0773****	(0.0158)				
Prom. sender and Prom. receiver[a]	-1.239	(1.114)			-3.984	(8.663)		
Prom. sender and St. low receiver	-2.002**	(0.941)			-0.514	(8.490)		
St. low sender and Prom. receiver	-0.510	(0.502)			-1.755	(3.436)		
St. high sender and Prom. receiver			0.0612	(0.258)			0.636	(1.583)
Dem. sender and Prom. receiver			-0.403	(1.037)			-11.56	(7.438)
Dem. sender and St. low trustee			-0.552	(1.070)			-9.255	(7.484)
Transfer in the DG, same school	0.356*	(0.191)	-0.203	(0.236)	1.590	(1.092)	0.733	(1.914)
Transfer in the DG, other school	0.239	(0.228)	0.583*	(0.315)	3.676**	(1.654)	6.338***	(2.316)
Female	-0.665	(0.747)	1.965**	(0.770)	-5.480	(5.424)	12.60**	(6.392)
Math quiz performance	0.00781	(0.259)	0.288	(0.175)	-0.111	(1.993)	-0.902	(1.292)
Risk attitude[b]	0.0804	(0.158)	-0.197	(0.198)	-0.0383	(1.403)	-1.470	(1.292)
Perception Math[a]	-0.227	(0.863)	2.198*	(1.132)	2.626	(8.403)	-9.159	(7.672)
Session F.E.	YES		YES		YES		YES	
N	174		174		174		174	
Pseudo R^2	0.114		0.227		0.068		0.137	
Log pseudolikelihood	-368.997		-293.796		-552.113		-487.819	
F	4.38		6.64		3.74		6.28	
$p > F$	0.000		0.000		0.000		0.000	

Notes: Marginal effects are reported. Robust standard errors clustered at the individual level are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. In models (1) and (2), the dependent variable is trust. Model (1) restricts observations to matches in which both the trustor and the trustee are from the business school. The reference category corresponds to steady-low trustors matched with a steady-low trustee. Model (2) restricts observations to matches in which the trustor is from the engineering school and the trustee is from the business school. The reference category corresponds to steady-high trustors matched with a steady-low trustee. In models (3) and (4), the dependent variable is trustworthiness. Model (3) restricts observations to matches in which both the truster and trustee are from the business school. The reference category corresponds to steady-low trustees matched with a steady-low trustor. Model (4) restricts observations to matches in which the trustee is from the engineering school and the trustor is from the business school. The reference category corresponds to steady-high trustees matched with a steady-low trustor. Prom. for promoted (business school), St. low for steady-low (business school), St. high for steady-high (engineering school), and Dem. for demoted (engineering school). (a)"Sender" refers to the trustor for the trust decision and to the trustee for the return decision. "Receiver" refers to the trustee for trust decision and the trustor for trustworthiness decision. (b) For risk attitudes, a greater number indicates less risk aversion. (c) Dummy indicating that the trustor chose "strongly disagree" or "disagree" about the statement *It is important to have good mathematical skills.*

Table A7.4 – Trust and Trustworthiness Toward Promoted Counterparts in the one school condition

	(1) Trust		(2) Trustworthiness	
Beliefs	0.140****	(0.029)		
Prom. sender and Prom. receiver	0.660	(1.910)	-3.485	(9.399)
Prom. sender and St. low receiver	0.437	(1.749)	0.979	(8.946)
St. low sender and Prom. receiver	0.137	(0.562)	-0.900	(1.997)
Transfer in the DG, same school	0.391*	(0.210)	2.411	(1.485)
Female	-0.193	(1.069)	2.909	(8.909)
Math quiz performance	-0.165	(0.492)	0.941	(2.723)
Risk attitude	0.308	(0.231)	0.870	(1.925)
Perception Math	1.506	(2.789)	-3.642	(7.847)
Session F.E.	YES		YES	
N	220		220	
Pseudo R^2	0.126		0.034	
Log pseudolikelihood	-428.315		-736.335	
F	5.11		2.43	
$p > F$	0.000		0.007	

Notes: Marginal effects are reported. Robust standard errors clustered at the individual level are in parentheses. ** $p < 0.05$.

In model (1), sender refers to trustor and receiver refers to trustee. In model (2) sender refers to trustee and receiver to trustor. The reference category corresponds to a steady-low matched with a steady-low . Prom. for promoted (business school), St. low for steady-low (business school), steady-high for steady-high (engineering school), and Dem. for demoted (engineering school). (a) For risk attitudes, a higher number indicates less risk aversion. (b) Dummy indicating that the trustor chose "strongly disagree" or "disagree" about the statement *"It is important to have good mathematical skills"*.

Chapter 2

Image spillovers in groups and misreporting.[1]

1 Introduction

Information about individuals is imperfect, and individuals are often judged as members of social groups rather than as individuals (Phelps, 1972; Arrow, 1973; Aigner and Cain, 1977). When it is the case, the image of the group is substituted for the less available individual information and determines the members' economic success. Some social groups have a positive image, which benefits the members. For instance, religious people tend to be perceived as more trustworthy (Tan and Vogel, 2008). On the other edge of the spectrum, some groups have a negative image. In these cases, the members may suffer from the negative stereotypes attached to the group and be excluded from profitable economic interactions. For instance, in an experiment with Israeli participants, Fershtman and Gneezy (2001) show that eastern Jews are believed to be less trustworthy than Ashkenazic Jews (which turns out to be erroneous). Similarly, Cohn et al. (2014) show that bankers are perceived to be less honest than people from the general population and about as honest as inmates. Even more problematically, a bad image at the

[1] Joint work with Daniel Houser.

group level can be erroneous and self-fulfilling (Glover et al., 2017).

Because the image of social group is a major determinant of the members' economic opportunities and success (e.g. Fershtman and Gneezy, 2001; Tan and Vogel, 2008; Doleac and Stein, 2013) understanding the way group members contribute to build it is important. It is clear that preserving a good individual image is an important driver of individual behavior: when exposed, people are more altruistic (e.g Andreoni and Bernheim, 2009), reciprocal (e.g. Bao et al., 2018) and honest (e.g. Kajackaite et al., 2018; Abeler et al., 2018). Social exposure increases norm-obedience because people want to preserve a good image for hedonistic and strategic reasons.

An important open question is whether this translates to situations in which one's behavior impacts the image of the group rather than the image of the individual. The image of social groups have been studied mainly through the lens of stereotypical thinking (e.g. Fershtman and Gneezy, 2001; Castillo and Petrie, 2010; Bordalo et al., 2016). In this literature, groups generally reflect an image that is rather exogenous. Only a few papers examine how concerns for the image of the group impacts the behavior of its members. Tirole (1996) shows theoretically that groups are confronted with a public good problem when it comes to building a collective reputation. Huck and Lünser (2010) show experimentally that groups can sustain a somewhat correct reputation: members forego selfish gains in order to preserve the image of the group. Relatedly, Eriksson et al. (2017) show that individuals are also willing to forego monetary gains to save an ingroup member from loosing face, even when group identity is minimal.

The present paper explores this question and identifies a potential mechanism of production for the social image of group. The premise of this mechanism is that, within groups, individuals share some underlying characteristics because of influence and selection. As a consequence, an external observer may use the first-hand information he has on one member of a group in order to decide how to interact with another member

of the group. We term this mechanism “image spillovers": the behavior of the observed one might spill over to the image of the the other one and (in part) determine his economic opportunities. Just like individuals are concerned about their individual image, members of groups might want to preserve the image of the group because they like to be in a group with a good image and because the image of the group determines the economic opportunities of the members. Consequently, anticipating that they might impact the image of the group could push them to behave better.

We study experimentally this mechanism in a setting in which the behavior of an individual might harm his ingroups' image, and, as a result, their chances of success in further interactions. More specifically, we test experimentally whether image spillovers in pairs of participants deter misreporting. The structure of the game, which is common knowledge, is the following. First, both members of the pair have to report the outcome of a random variable independently and without communication. They can report truthfully or misreport. Generally, misreporting pays more. Second, an external observer (the Audience), observes the report of one predetermined member of the pair, and then has to decide whether he is willing to interact with the other member of the pair. The Audience wants to play only if he expects that the probability of a truthful report by the second member is high enough. Both members of the pair prefer the Audience to play, because when the Audience plays, they get an extra payoff.

We implement two experimental conditions, between which we vary whether image spillovers can arise across the members of the pair. In the baseline, in each pair, members of the pair do not know each other and were invited independently. In the treatment, the members of each pair are friends, and were invited jointly. The idea behind this treatment is that friends tend to be similar. This view is widely supported in economics (e.g. Leider et al., 2009; Goeree et al., 2010) as well as other social sciences (e.g. Kandel, 1978; McPherson et al., 2001; Lee et al., 2009) and even biology (e.g. Christakis and Fowler, 2014). This reflects both homophily (the tendency to become friends with similar

individuals) and influence (the tendency of friends to exert influence on each others). If the Audience is aware of this, friends share a common image: the Audience can pull information from the behavior of the first friend and condition his decision to play with the second friend on this information. We do not expect such a pattern in the baseline. In turns, if friends whose reports are observed anticipate the impact of their reports on the decision of the Audience, they should misreport less often in the treatment than in the baseline in order to try and convince the Audience to play. To sum up, we expect image spillovers to arise in the treatment but not in the baseline, and the perspective of image spillovers to deter misreporting.

We find that image spillovers do not deter misreporting: participants whose reports are observed are equally likely to misreport in both conditions. However, in the treatment,the Audiences use the observation of the first members of the pairs to decide whether to play or not. This behavior corresponds to the belief that friends are more likely to take the same decision than two participants taken randomly. These beliefs are in turn corroborated by the data. By analyzing the participants' beliefs, we discuss potential explanations for this. We argue that participants whose reports are observed typically underestimate the impact of their behavior on the decision of the Audience.

While specific, we consider studying image spillovers in pairs of friends as a good foray in the study of image spillovers in groups in general. Indeed, in our experiment, group membership is salient for the members of the pair, and the lack of effect of image spillovers on misreporting cannot be blamed on participants ignoring they are part of a group. Moreover, the mechanism we are interested in rely on the fact that members of social groups share a common image because they share (or are believed to share) common characteristics, which can influence both how non-group members interact with ingroups and the behavior of ingroups themselves. This is a characteristic that pairs of friends share with a wider range of groups (e.g. Alma Matter, religious group,

profession).

The remaining of the paper is organized as follows. In section 2, we present the experimental design. In section 3, we introduce the conceptual framework and the hypotheses to be tested. In section 4, we present our results. Last, section 5 concludes.

2 Experimental design

2.1 The experimental game

The experimental game aims at capturing the impact a group member might have on his ingroups' economic opportunities through the impact of his own behavior on the image of his group.

The game is played for four periods. Repetitions were necessary to be able to perform within-subject statistical tests. Before the first period, the program assigns their roles to participants. They are either Audience or member of a pair. In each pair, one participant is given the role of f_1 and the other is given the role of f_2. Pairs and roles remain fixed for the length of the experiment.

For each pair, the program draws a card among Jack, Queen, King, Ace and Joker (all diamonds). This card is called the pair's card. It is common knowledge that the pair's card remains the same for the length of the experiment. In each pair, f_1 and f_2 are informed of the pair's card.

Across periods, each pair is matched to a different Audience and each pair is matched to each Audience exactly once. To avoid learning, no feedback about the outcome of the game is given between periods.[2] At the end of the session, one period is randomly

[2]This is especially important for our analysis of the behavior of the Audience, which mainly relies on within-subject tests. Giving no information between periods reduces the possibility of dynamic session-effects Fréchette (2012).

selected to determine the payoff of every participant in the session. Participants are informed of their payment for the selected period only, and they are not informed of the payment of any other player. This insures that, within a pair, one participant cannot infer the decision of the other from his payment.

At the beginning of each period, the program draws a wining card for each pair independently, among Jack, Queen, King, Ace or Joker. This card can be different from the pair's card. f_1 and f_2 are informed of the winning card on their screen. Then, f_1 and f_2 are instructed to report their pair's card simultaneously and without any communication allowed. However, they can chose to report any card they want among Jack, Queen, King, Ace or Joker. Not reporting the card of the pair is cheating since it violates the rule. Participants are informed that their report is private, except for the Audience with whom they are matched for the present period.

After f_1 and f_2 have made their report, the Audience observes the card of the pair and the report of f_1, but not the report of f_2. Thus, the Audience is able to see whether f_1 misreported, but not if f_2 misreported. The Audience then decides whether to "pass" or "play".

f_1's payoff is the sum of two elements. If f_1 reports the winning card, he earns 1125 ECU (125 ECU = €1). If f_1 reports another card, he earns 875 ECU. The second element depends on the decision of the Audience: if the Audience decides to play, f_1 earns an extra 625 ECU.

f_2's payoff depends on the decision of the Audience: f_2's decision matters if and only if the Audience plays. If the Audience does not play, f_2 earns a fixed amount of 750 ECU. If the Audience plays, f_2 earns 1375 ECU when reporting the winning card, and 1000 ECU when reporting another card. Thus, f_1 and f_2 are always better off when the

Audience plays.[3]

If the Audience passes, his payoff is 875 ECU. If the Audience plays, his payoff depends on the report of f_2: if f_2 reported truthfully, the Audience earns 1250 ECU. If f_2 misreported, the Audience earns 625 ECU.[4]

2.2 Experimental treatments

Between our two conditions, we vary the potential for image spillovers. The only difference between the baseline and the treatment is that the members of each pair in the treatment are friends in the real life. This is not the case in the baseline.

In practice, we recruited couple of friends for both conditions. In the invitation email, it was made clear that participants should come to the lab with a friend, otherwise they would not be allowed to participate. In the treatment, the matching protocol was such that two friends were either both Audience or both members of the same pair. In the baseline, the matching protocol insured that two friends could never interact: two friends would either be both Audience or members of two different pairs. Participants were informed about the matching protocol in the instructions. Moreover, in the instructions, members of pairs were called Friend 1 and Friend 2 in the treatment, and Person 1 and Person 2 in the baseline (see Appendix A1).

[3]The payoff functions of f_1 and f_2 insure that f_1 always gets a higher payoff than f_2, whatever the situation. This evacuates ranking considerations: f_1 cannot be motivated to misreport in order to be sure to be paid more than his friend. Such motivations have been shown to be strong driver of behaviors (see e.g. Charness et al., 2014).

[4]In addition, after the fourth period of the game, participants played a mind coin toss game. The program simulated a random draw and participants had to report whether they predicted the outcome correctly. The aim was to test whether, in a pair of friends, the behavior of one player is correlated with the behavior of the other player in a context in which payoffs are independent. Since only 16,5% of the participants reported that they did not predict the outcome correctly, we do not use these data in the remaining of the paper. The data is available upon request.

2.3 Post experimental questionnaire

At the end of the session, but before any information about payoffs, participants had to make five guesses about how participants behaved in the session. In each of the five scenarios, there were two possible decisions. Participants were asked to predict the proportion of each decisions in their session. For instance, they had to predict the proportion of truthful reporting (resp. misreporting) by f_2s. Answers were incentivized: at the end of the session, one prediction was randomly drawn. The closer this prediction was from the true proportion, the more money they could earn.[5]

The first scenario corresponds to the decision made by f_1: participants had to predict the proportion of misreporting by f_1s. The next two scenarios correspond to the decision made by the Audience. Participants had to guess the proportion of decisions to play conditional on the report by f_1. First, they had to guess the proportion of decision to play after observing truthful reporting. Then, they had to guess the proportion of decisions to play after observing misreporting. The last two scenarios correspond to the decision made by f_2. First, they had to guess the proportion of decisions to misreport knowing that f_1 misreported. Second, they had to guess the proportion of decisions to misreport knowing that f_1 reported truthfully. If a participant expects that the decisions of f_1 and f_2 are independent, the two guesses about f_2 should not differ. Conversely, if participants expect that for some reasons these decisions are not independent, the predictions should differ.

Then, participants had to fill out a questionnaire aimed at checking that friends who came together to the sessions were actually friends. To asses the degree of friendship of participant, we asked the same question to each participant about himself and about

[5]The full payoff function is as follows : If the prediction was equal to the actual value more or less 5 percentage point, they earned 250 ECU. If the prediction was between 5 and 10 percentage point away from the actual value, they earned 187.5 ECU. If the prediction was between 10 and 15 percentage point from the actual value, they earned 125 ECU. If the prediction was between 15 and 20 percentage point from the actual value, they earned 62 ECU. If the prediction was further than 20 percentage points from the actual value, they earned nothing.

his friend (for instance “Do(es) you (your friend) have a pet?") and compared answers to measure mutual knowledge. We also used a validated questionnaire by Mendelson and Aboud (1999) that measures the quality of friendship relations. Data from these questionnaires confirms that participants who came together were likely real-life friends (see appendix A1 for the details of these questionnaires and data supporting our claim).

Finally, we elicited socio-demographic characteristics, including a measure of risk attitudes using the procedure of Dohmen et al. (2011).

2.4 Experimental Procedures

The experiment was developed using z-Tree (Fischbacher, 2007). All sessions were conducted at GATE-LAB, Lyon. There were 10 sessions in total (4 sessions for the baseline, 6 for the treatment). For each session, we recruited 6 couples of friends. In total, 120 participants took part in the experiment. Participants were recruited using Hroot (Bock et al., 2014). In the recruitment email, we explained to the potential participants that they had to come with a friend to the laboratory in order to be allowed to take part in the session.

Participants were randomly assigned to individual cubicles. Friends were never assigned to adjacent cubicles. The instructions were distributed to participants and read aloud by the same experimenter (see Appendix A1).

Before the first period, participants had to fill out a comprehension questionnaire. They could not proceed to the experiment before completing the questionnaire correctly. Questions were answered in private.

The average length of a session was 50 minutes. The average payoff was €14.78 (Min: €10, Max: €22, standard deviation, S.D. hereafter: 2.85). Payments were made in cash, in a separate room and in private.

3 Framework and hypotheses

To formalize the effect of the treatment on misreporting, we take the point of view of a participant i in the role of f_1 who has to decide whether to misreport or report truthfully. Equation 3.3 represents his expected utility.

$$\mathbb{E}u_i(x, e, \theta_i) = \pi(x) - c(x, \theta_i) + p(x, e)M(e) \tag{2.1}$$

x corresponds to his decision, $x = m$ if he misreports and $x = h$ if he reports truthfully. e represents the experimental condition, ($e = b$ for the baseline and $e = t$ for the treatment).

$\pi(x)$ represent the direct consequences of the participant's report on his own monetary payoff. Participants can earn more money by misreporting if the winning card differs from the pair's card: $\pi(m) > \pi(h)$. $c(x, \theta_i)$ is the moral cost of misreporting. We set $c(h, .) = 0$ and consider that $\partial c(m, \theta_i)/\partial \theta_i > 0$. This gives a direct interpretation to θ_i, as an idiosyncratic preference parameter for truthful reporting. It captures the fact that participants are heterogeneous in this respect.[6] $c(x, \theta_i)$ also accounts for the potential affective value of having a good image in the eyes of other participants and the experimenter.

$p(x, e)M(e)$ formalizes the mechanism of image spillovers. $p(x, e)$ represents the participant's beliefs on the decision of the Audience conditional on his report x and the experimental condition e. We define $\Delta(e) = p(h, e) - p(m, e)$ as the impact of moving from misreporting to truthful reporting in condition e, as expected by f_1. We consider that $\Delta(e) \geq 0$. This means that the Audience is expected to be at least as likely to play after observing truthful report than after observing misreporting. We further assume that $\Delta(t) > \Delta(b)$. This captures the fact that participants expect the Audiences to condition their decisions on the decisions of f_1 to a greater extent in the treatment than in the

[6] For instance, Gibson et al. (2013) find an important heterogeneity in preference for truthfulness.

baseline, because they expect the Audience to believe that friends tend to be similar. We should stress that this does not imply that the Audiences does not condition their decisions in the baseline.[7]

$M(e)$ represents the extra utility for participant i when the Audience plays. $M(e)$ is composed of a monetary part and a non-monetary part. The monetary part accounts for the extra payoff for f_1 when the Audience decides to play. It is constant across treatment, such that $M(e) > 0$. The non-monetary part accounts for the fact that f_1 potentially cares about f_2's payoff, which is greater when the Audience plays. Assuming that f_1 cares more about his friend than about another participant, $M(t) > M(b)$.

f_1 is indifferent between misreporting and reporting truthfully if his type $\hat{\theta}$ is such that equation 2.2 holds.

$$\mathbb{E}u_i(m, e, \hat{\theta}) = \mathbb{E}u_i(t, e, \hat{\theta}) \tag{2.2}$$

It is straightforward to show that for $\theta_i < \hat{\theta}$, f_1 misreports and for $\theta_i > \hat{\theta}$, f_1 reports truthfully. Equation 2.2 can be rearranged to give equation 2.3.

$$c(x, \hat{\theta}(e)) = \pi(m) - \pi(t) - \Delta(e)M(e) \tag{2.3}$$

$\hat{\theta}(e)$ is the type that makes participants indifferent between misreporting and reporting truthfully in condition e. Our comparative static exercise consists in comparing $\hat{\theta}(b)$ and $\hat{\theta}(t)$. Equation 2.4 summarizes how the limit type is expected to change between experimental conditions.

$$c(x, \hat{\theta}(b)) - c(x, \hat{\theta}(t)) = \Delta(t)M(t) - \Delta(b)M(b) \tag{2.4}$$

[7]There are several reasons that could lead Audience to condition his decision in the baseline. For instance, observing misreporting could prime the Audience with the idea that misreporting is a widespread behavior.

Equation 2.4 implies that $\hat{\theta}(b) > \hat{\theta}(t)$: because image spillovers can occur in the treatment, some individuals who would misreport in the baseline report truthfully in the treatment. This is because his decision is expected to have a greater impact on the Audience's decision in the treatment compared to the baseline ($\Delta(t) > \Delta(b)$), and because the gain if Audience plays is at least as large in the treatment compared to the baseline ($M(t) \geq M(b)$). This formalizes the (expected) deterrent effect of image spillovers.

From the previous framework, we derive three hypotheses.

Hypothesis 1: The behavior of f_1 is predictive of the behavior of f_2 in the treatment, not in the baseline. This hypothesis corresponds to testing whether friends are "similar". Previous literature indicates that friends cluster by preference because of social influence and homophily (e.g. McPherson et al., 2001; Leider et al., 2009; Goeree et al., 2010), which should translate into behavior regarding rule breaking.

Hypothesis 2: In the treatment, the Audience plays less frequently after observing misreporting than after observing truthful reporting. No such difference is expected in the baseline. This hypothesis illustrates the fact that observing bad behavior on the behalf of a member of a social group harms the opportunity of other group members. We further conjecture that the effect is driven by *beliefs updating*: when f_1 and f_2 are friends, the Audience updates her beliefs about f_2 upon observing how f_1 behaved. An alternative mechanism would be that the Audience wants to punish individuals who misreported, in the spirit of third party punishment, but by design, we made third party punishment very costly for the Audience and thus not likely to be a driver of behavior. Moreover, this motivation is in principle constant between experimental conditions.

Hypothesis 3: There is less misreporting by f_1 in the treatment compared to the baseline. This hypothesis corresponds to the deterrent effect of image spillovers. Consistently with the framework presented above, we expect that some participants who

would misreport in the baseline decide to report truthfully in the treatment. We expect them to internalize the harm their behavior can have on f_2's image in the treatment.

4 Results

Table 4.1 summarizes the main variable of interest, namely the frequency of misreporting by f_1 and f_2, and the decision to play by the Audience.[8]

Table 4.1 – Summary statistics.

Condition	Baseline			Treatment			Pooled		
Variable	Mean	Std. Dev.	N	Mean	Std. Dev.	N	Mean	Std. Dev.	N
f_1 misreports	0.615	0.491	52	0.584	0.496	77	0.596	0.492	129
f_2 misreports	0.673	0.473	52	0.649	0.480	77	0.658	0.475	129
Audience plays	0.403	0.495	52	0.441	0.499	77	0.42	0.496	129

f_1 (f_2) misreport=1 if f_1 (f_2) reported a card that was not his pair's card
Audience plays=1 if the Audience plays

Result 1: The behavior of pair members is correlated only when pairs are composed of two friends. This result supports Hypothesis 1.

Support for result 1: Figure 4.1 represents the proportion of misreporting by f_2 for each report of f_1 by condition. It suggests that the behavior of f_1 and f_2 are correlated in the treatment exclusively. In the baseline, the frequency of misreport by f_2 goes from 60% when f_1 reported truthfully to 71.9% when f_1 misreported (Fischer exact test: $p = 0.544$). In the treatment, it goes from 46.9% when f_1 reported truthfully to $77,6\%$ when f_1 misreported (Fischer exact test: $p = 0.008$).

[8]Unless otherwise stated, we drop observations in which subjects could not increase their payoff by misreporting, i.e. observations for which the winning card is the same as the pair's card: in this situation, f_1 and f_2 are always better-off reporting truthfully, and Audiences are always better-off playing. In these cases, participants played the dominant strategy 84.6% of the times. 9% of the participants deviated from the dominant strategies at least once (10.4% in the baseline, 8.3% in the treatment). Excluding them from our analysis does not change our results.

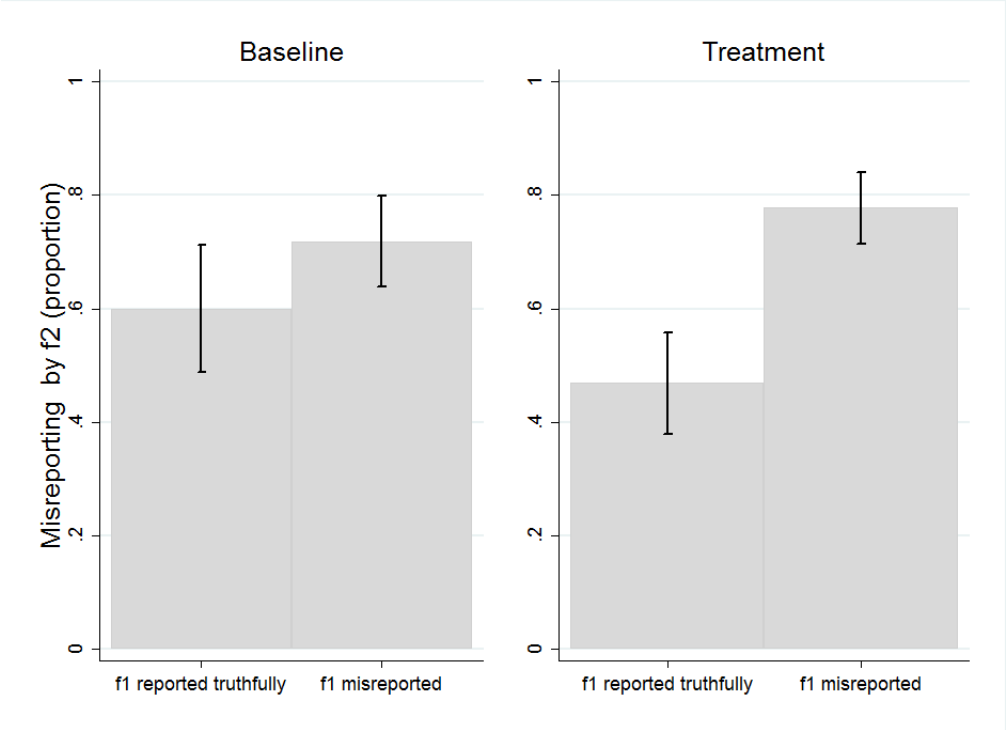


Figure 4.1 – Proportion of misreporting by f_2 conditional on f_1's report, over condition.

Note: The vertical lines represent standard errors.

In order to account for multiple observations, we test whether decisions by f_1 and f_2 are more likely to match than if they were totally independent. First, we compute the theoretical (benchmark) proportion of matched decision for each pairs. We sum the expected proportion of the event "both f_1 and f_2 misreported" and "both f_1 and f_2 reported truthfully". In each case, the expected proportion is computed as the product of the proportion of misreport (resp. truthful report) by f_1 with the proportion of misreport (resp. truthful report) by f_2. Second, we compute the empirical frequency of matched decisions for each pair of f_1 and f_2. We sum the empirical frequency of the event "both f_1 and f_2 misreported" and "both f_1 and f_2 reported truthfully". Then, for each pair, we compare the empirical frequency to the theoretical proportion using a Wilcoxon matched-pairs signed-ranks test (W hereafter). Under the null hypothesis, the empirical

frequency is not different from the theoretical one. Under the alternative, the empirical frequency is greater than the theoretical one.[9] Results are reported in Table 4.2.

Table 4.2 – Comparison of empirical and theoretical frequencies of match between f_1 and f_2 decisions

	(1) Empirical frequency	(2) Theoretical frequency	(3) Diff	(4) N	(5) $p-$value (1-sided)
Baseline	0.614	0.539	0.075	16	0.189
Treatment	0.667	0.525	0.142	24	0.047
Pooled	0.645	0.53	0.115	40	0.028

Column (1) reports the empirical frequency of match between decisions of f_1 and f_2 (either both misreported or both reported truthfully. Column (2) reports the theoretical frequency, i.e. the frequency of matches expected if the decisions of f_1 and f_2 were independent. Column (3) reports the difference between column (1) and (2). Column (5) reports the $p - values$ for one-sided Wilcoxon matched-pairs signed-ranks test that the empirical frequency is greater than the theoretical one.

Overall, the frequency of matching decisions is greater than what is expected under the null (difference$=$ 0.115, $p = 0.028$). In the treatment, f_1's and f_2's decision match significantly more often than what is expected under the null (difference$=$ 0.142, $p = 0.047$). In the baseline, the difference between the empirical frequency and the theoretical frequency is not statistically significant (difference$=$ 0.075, $p = 0.189$).

Result 2: In the treatment, the Audience plays less often when f_1 misreports. This is not the case in the baseline. This result supports Hypothesis 2.

Support for result 2: Figure 4.2 depicts the raw figures of decisions to play, conditional on the report by f_1, by condition.

[9] Our test is one sided because our hypothesis follows the bulk of the literature on friendship that emphasizes that friends tend to be alike, rather than different.

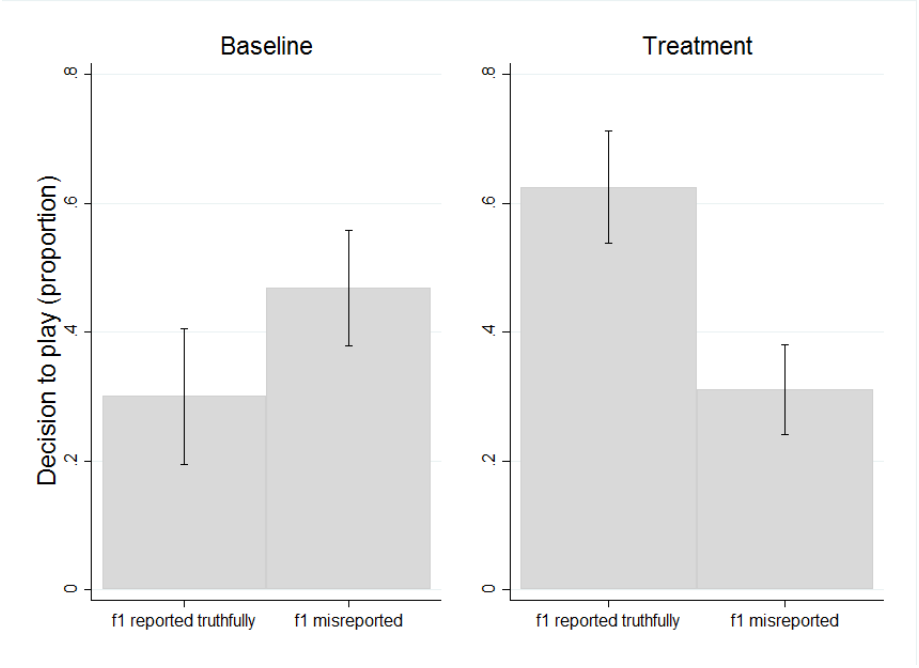


Figure 4.2 – Play decision conditional on report by f_1s, by condition
Note: The vertical lines represent standard errors.

In the baseline, Audience played 30% of the time after observing a truthful report and 46.9% after observing misreporting. In the treatment, Audience played 62.5% of the time after observing a truthful report and 31.1% after observing a untruthful report.

In order to formally test that the Audiences play less often when f_1 misreports than when f_1 reported trutfhully, we compute for each Audience the frequency of decisions to play after observing misreporting and the frequency of decisions to play after observing a truthful report. We compare these frequencies within Audience. Under the null hypothesis, the frequency of decisions to play after observing misreporting is not different from the frequency to play after observing a truthful report.

The difference between the frequency of decisions to play after observing truthful report and after observing misreport is not significant in the baseline (W: $p = 0.618$)[10] but is significant in the treatment (W: $p = 0.016$). This suggests that Audiences are conscious of the correlation between the behavior of f_1 and f_2, when f_1 and f_2 are friends. Table A2.1 in Appendix A2 confirms with logit regressions that this result is robust to the introduction of individual characteristics and session fixed-effects.

The theoretical motivation for the hypothesis supported by Result 2 is that the Audience can pull more information from the behavior of f_1s in the treatment than in the baseline. In order to explore this mechanism, we measure the extent to which beliefs about f_2 is conditional on f_1's report. For each Audience, we compute the difference between guesses when f_1 reported truthfully and when f_2 misreported. Audience guessed that f_2s were 4.75% less likely to misreport when f_1 misreported in the baseline, but 9% more likely in the treatment. The previous comparison is qualitatively consistent with our expectations but does not reach significance (Two-sided Mann-Whitney test, MW hereafter, $p = 0.13$).

A complementary analysis is to test if those who guessed correctly, i.e. those who, in the treatment, guessed that f_2 would misreport more when the friend had misreported in the first place, are those who renounced to play after observing misreporting. To do so, we estimated logit models which explain the decision to play of the Audience by the Audience's guesses conditional on observing misreporting. The explanatory variables are a dummy variable indicating that f_1 misreported interacted with the variable "update" that proxies the extent to which the Audience updates his beliefs about f_2 after observing f_1. The variable "update" is the difference between the guess about the probability that f_2 misreports after observing a truthful report and after observing a misreport. When the variable "update" is positive, the participants guessed that it is more likely that f_2

[10] In a one-sided Wilcoxon matched-pairs signed-ranks test, for a power of 0.8 and an Alpha-risk of 0.05, our design allows for a detection of a medium sized effect according to the standard classification (Cohen, 1988).

will misreport after observing a misreport by f_1 than after observing a truthful report. We also control for the belief of the Audience about the propensity of f_2 to misreport upon observing a truthful report by f_1. In all specifications, we control for risk attitudes, gender, age, monthly expenses and whether participants have experience with economic experiments. We also include session fixed effects. Table 4.3 reports the marginal effect of the variable "update" conditional on observing misreporting.

Overall, the variable "update" has a significant effect on the Audience's decision to play conditional on f_1 misreporting (column (1), $p = 0.045$). Breaking down by experimental condition, the effect is not statistically significant in the baseline (column (2)), but is larger and highly statistically significant in the treatment (column (3)) ($p < 0.001$). This suggests that the mechanism underlying result 2 relies on updating : Audiences pull more information from the observation of f_1, and use it more, when f_1 and f_2 are friends.

Table 4.3 – The effect of updating on the decision to play conditional on observing misreporting by f_1

	(1)	(2)	(3)
	$play = 1$	$play = 1$	$play = 1$
update	-0.004**	-0.003	-0.011****
	(0.002)	(0.004)	(0.003)
Session F.E.	Yes	Yes	Yes
Ind. char.	Yes	Yes	Yes
N	129	52	77
Clusters	40	16	24

Marginal effect reported. Robust standard errors clustered at the individual level in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$

Since the behavior of f_1 is predictive of the behavior of f_2 in the treatment and the Audiences use this information, it is natural to test whether Audiences make "the right decision" more often in the treatment compared to the baseline. We consider that right decisions for the Audience are decisions that maximizes her payoff ex-post. There are two right decisions: (i) to play when f_2 reported truthfully and (ii) no to play when

f_2 misreported. In order to assess whether Audiences make the right decision in the treatment more often , we computed the rate of right decisions for each individual Audience, and compared these rates between conditions. On average, Audiences make optimal decisions 48.43% of the times in the baseline and 61.15% of the times in the treatment (MW, two-sided $p = 0.065$). This is the good news about image spillovers: external observers can pull valuable portable information from the observation of the behavior of ingroups.

Result 3: Image spillovers do not foster truthful reporting. This result goes against hypothesis 3. Observed players do not adjust their behavior to the anticipation that their decision will influence the Audience's decisions.

Support for result 3: f_1 misreported $61,5\%$ of the time in the baseline and $58,4\%$ of the time in the treatment (misreporting rate are averaged at the individual level, Mann-Whitney test, $p = 0.632$).[11] In Table A2.2 in Appendix A2, we present the result of logistic regressions in which misreporting by f_1 is explained by the experimental condition. Results from these regressions confirm the absence of effect of image spillover on misreporting.

Result 3 sets off two potential explanations. The first one relates to preferences. f_1s could consider that it is not worthwhile to refrain from misreporting even if they expect that their behavior may impact the Audiences' decisions: misreporting yields a certain payoff (250 ECU, €2) whereas truthful reporting has an uncertain impact on payoff, which depends on the Audience's decision. This explanation suffers the following objection. If the Audience decides to play, it yields an amount of money to f_1 that is considerably more important than misreporting (625 ECU for playing *vs.* 250 ECU for misreporting). In the treatment, more risk tolerant individuals should refrain from misreporting more often to have a chance to convince the Audience to play. We should

[11] In a two-sided Mann-Whitney test, for a power of 0.8 and an Alpha-risk of 0.05, our design allows for a detection of a medium sized effect according to the standard classification (Cohen, 1988).

thus see a relation between misreporting and risk tolerance. Spearman's correlation coefficient between individual risk tolerance and individual misreporting frequency are very low and never significant if we pool all observations ($\rho = -0.02$, $p = 0.898$), or if we separate by condition (baseline : $\rho = 0.088$, $p = 0.745$, treatment : $\rho = -0.103$, $p = 0.631$).

The second potential explanation relates to the perception of f_1s about the impact of their report on the decisions of the Audiences. As shown earlier, in the treatment, the Audiences played 61.5% of the time when observing truthful reporting and 31.1% of the times when observing misreporting, which represents a drop of 31.4% percentage point. f_1 guessed that Audiences would play 57% of the time when observing truthful reporting and 50.16% of the time after observing misreporting. This difference is significantly different from zero (W, $p = 0.057$), but also significantly lower than the actual impact of observing misreporting on the decision to play (W, $p < 0.001$). f_1s underestimate the impact of their report on the decision of the Audience. This in turn can explain why f_1s misreport as often in the treatment and in the baseline.

A possible source for these mispecified beliefs is that participants in the role f_1 fail to realize that the behavior of f_1s predicts the behavior of f_2s in the treatment. Interestingly, f_1s are more likely than participants in other roles to guess that f_2s would report truthfully more often when f_1 had misreported than when f_1 had reported truthfully (Fischer exact test: $p = 0.024$). An interpretation is that assignment to role f_1 determined beliefs of participants in a self-serving manner, although belief elicitation was incentivized.[12] Believing that friends are actually different allows f_1s to hope that Audiences would play even after observing misreporting.[13]

[12]The impact of role assignment on beliefs has been studied before: Babcock et al. (1995) show in a dispute settlement experiment that what participants deem fair depends on their role in the experiments.

[13]In our experiment, participants in the role of f_1 had no opportunities to learn across periods, which is a potential explanation for the mispecified beliefs. Nevertheless, participants in other roles had no opportunities to learn either but do not hold such mispecified belief.

A consequence of the aforementioned mispecified beliefs is found in how misreporting by $f1$s impacts f_1s' and f_2s' payoffs. In the baseline, f_1s have greater payoffs per period on average when they misreport (1417.9 ECU when f_1 misreports and 1062.5 when they reported truthfully, MW $p < 0.001$), but this is not the case in the baseline (1265.6 *vs.* 1291.6, MW $p = 0.50$). Because misreporting deters the Audience from playing, it is not profitable for f_1 to misreport. Moreover, in the treatment, misreporting by f_1s reduces f_2's profit (927.7 *vs.* 1000, MW $p = 0.06$). Taken together, this indicates that, probably because $f1$s fail to anticipate the consequences of misreporting on their friends' image, they act in a way that is at least contrary to their friends interest, and probably against their own interest, given that f_1s may care about their friends.

5 Conclusion

In this paper, we study how group members contribute to the image of the ingroup. In an experiment with pairs of friends, we tested whether individuals internalize the fact that they can harm their ingroup's opportunities by misbehaving if an observer infers a correlation between ingroups' decisions.

We find no evidence of such a mechanism. While external observers used the information contained in one's friend's behavior to update their beliefs about the other friend, participants did not internalize the impact of their behavior on their friend's image. This is the case although they have monetary incentives to preserve the image of the pair of friends they belong to. Our preferred interpretation is that participants strongly underestimate the harm they do to the economic opportunities of their friends. They expected their friend's economic opportunity to be only marginally affected by their behavior because they did not fully realize that they share a common image.

The results of the present paper contains a pessimistic message: even in situation in which a group member has substantial incentives to behave such as to preserve the

group's image, mere ignorance of the mechanism leading to the production of group image can lead to hurt the group's image. As a consequence, the well identified disciplining effect of exposure found at the individual level (e.g. Andreoni and Bernheim, 2009; Bao et al., 2018; Kajackaite et al., 2018; Abeler et al., 2018) does not necessarily translate to situations in which image is collective rather than private.

In a more general perspective, our results identify a possible mechanism for poor social image at the group level. Group members fail to internalize that their behavior will impact the other members' opportunities. In our experiment, it is in their monetary interest to preserve the group's image so that this failure is not due to the fact that they do not care about their ingroups' fate, but to the fact that they fail to see that, whatever they do, it might be considered as a signal on the behavior of the members of their group.

Our results also contain a more positive message. In our experiment, strategical considerations related to the image of the group did not push dishonest participants to mimic honest participants. Thus, external observers were able to use the information contained in one's behavior to update their beliefs. Consequently, external observers took more profitable decisions in interacting with ingroups.

Several interesting related questions could be investigated. For instance, we could study image spillovers in the context of benevolent actions. Would external observers still use the information of one group member to update his beliefs about other group members? Would group members internalize the positive impact that their benevolent actions have on the image of the group? We could also allow participants to learn that their action impacts the social image of their group by giving them feedback across rounds. Would it correct mispecified beliefs and help group members internalize the effect of their behavior on the image of the group?

Bibliography

Abeler, J., D. Nosenzo, and C. Raymond (2018). Preferences for truth-telling. *Econometrica*.

Aigner, D. J. and G. G. Cain (1977). Statistical theories of discrimination in labor markets. *ILR Review 30*(2), 175–187.

Andreoni, J. and B. D. Bernheim (2009). Social image and the 50-50 norm: A theoretical and experimental analysis of audience effects. *Econometrica 77*(5), 1607–1636.

Arrow, K. J. (1973). *The theory of discrimination*, pp. 3–33. Princeton University Press.

Babcock, L., G. Loewenstein, S. Issacharoff, and C. Camerer (1995). Biased judgments of fairness in bargaining. *The American Economic Review 85*(5), 1337–1343.

Bao, W., Y. Rao, J. Wang, and D. Houser (2018). Social exposure and trustworthiness: Experimental evidence. *Economics Letters 162*, 73 – 75.

Bock, O., I. Baetge, and A. Nicklisch (2014). hroot: Hamburg registration and organization online tool. *European Economic Review 71*, 117 – 120.

Bordalo, P., K. Coffman, N. Gennaioli, and A. Shleifer (2016). Stereotypes. *The Quarterly Journal of Economics 131*(4), 1753–1794.

Castillo, M. and R. Petrie (2010). Discrimination in the lab: Does information trump appearance? *Games and Economic Behavior 68*(1), 50 – 59.

Charness, G., D. Masclet, and M. C. Villeval (2014). The dark side of competition for status. *Management Science 60*(1), 38–55.

Christakis, N. A. and J. H. Fowler (2014). Friendship and natural selection. *Proceedings of the National Academy of Sciences 111*(Supplement 3), 10796–10801.

Cohen, J. (1988). Statistical power analysis for the behavioral sciences. 2nd.

Cohn, A., E. Fehr, and M. A. Maréchal (2014). Business culture and dishonesty in the banking industry. *Nature 516*(7529), 86.

Dohmen, T., A. Falk, D. Huffman, U. Sunde, J. Schupp, and G. G. Wagner (2011). individual risk attitudes: measurement, determinants, and behavioral consequences. *Journal of the European Economic Association 9*(3), 522–550.

Doleac, J. L. and L. C. Stein (2013). The visible hand: Race and online market outcomes. *The Economic Journal 123*(572), F469–F492.

Eriksson, T., L. Mao, and M. C. Villeval (2017, Sep). Saving face and group identity. *Experimental Economics 20*(3), 622–647.

Fershtman, C. and U. Gneezy (2001). Discrimination in a segmented society: An experimental approach. *The Quarterly Journal of Economics 116*(1), 351.

Fischbacher, U. (2007). z-tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics 10*(2), 171–178.

Fréchette, G. R. (2012, Sep). Session-effects in the laboratory. *Experimental Economics 15*(3), 485–498.

Gibson, R., C. Tanner, and A. F. Wagner (2013, February). Preferences for truthfulness: Heterogeneity among and within individuals. *American Economic Review 103*(1), 532–48.

Glover, D., A. Pallais, and W. Pariente (2017). Discrimination as a self-fulfilling prophecy: Evidence from french grocery stores. *The Quarterly Journal of Economics 132*(3), 1219–1260.

Goeree, J. K., M. A. McConnell, T. Mitchell, T. Tromp, and L. Yariv (2010, February). The 1/d law of giving. *American Economic Journal: Microeconomics 2*(1), 183–203.

Huck, S. and G. K. Lünser (2010). Group reputations: An experimental foray. *Journal of Economic Behavior & Organization 73*(2), 153–157.

Kajackaite, A., U. Gneezy, and J. Sobel (2018). Lying aversion and the size of the lie. *The American Economic Review*.

Kandel, D. B. (1978). Homophily, selection, and socialization in adolescent friendships. *American Journal of Sociology 84*(2), 427–436.

Lee, K., M. C. Ashton, J. A. Pozzebon, B. A. Visser, J. S. Bourdage, and B. Ogunfowora (2009). Similarity and assumed similarity in personality reports of well-acquainted persons. *Journal of personality and social psychology 96*(2), 460.

Leider, S., M. M. Möbius, T. Rosenblat, and Q.-A. Do (2009). Directed altruism and enforced reciprocity in social networks. *The Quarterly Journal of Economics 124*(4), 1815–1851.

McPherson, M., L. Smith-Lovin, and J. M. Cook (2001). Birds of a feather: Homophily in social networks. *Annual review of sociology 27*(1), 415–444.

Mendelson, M. J. and F. E. Aboud (1999). Measuring friendship quality in late adolescents and young adults: Mcgill friendship questionnaires. *Canadian Journal of Behavioral Science 31*(2), 130–132.

Phelps, E. S. (1972). The statistical theory of racism and sexism. *The american economic review 62*(4), 659–661.

Tan, J. H. and C. Vogel (2008). Religion and trust: An experimental study. *Journal of Economic Psychology 29*(6), 832–848.

Tirole, J. (1996). A theory of collective reputations (with applications to the persistence of corruption and to firm quality). *The Review of Economic Studies 63*(1), 1–22.

A1 Assessment of friendship

In order to make sure that we were successful in recruiting friends to the lab, we collected two measures of friendship. The first one is original and relies on a measure of mutual knowledge. The second one is drawn from the literature in social psychology and measures the quality and the satisfaction of a given friendship.

Measure of mutual knowledge

We designed a questionnaire that was aimed at measuring mutual knowledge. The underlying assumption is that the more two people (friends) interact, the better they know each other. It was implemented at the end of the session. Participants had to answer to a series of nine questions about themselves, and then they had to answer to the same questions about their friends. The questions are the following:

- The participant answers the following question about him
    1. Do you have a pet? (Yes, No)
    2. If you have a pet, what kind? (Open)
    3. If you have a pet, what is the name of the pet? (Open)
    4. Do you have any brother or sister? (Yes, No)
    5. If you have brother(s) and / or sister(s), how many? (Open)
    6. If you have brother(s) and / or sister(s), what is the name of the elder? (Open)
    7. Do you practice a sport? (Yes, No)
    8. If yes, what is this sport? (Open)
    9. If which season were you born? (Winter, Spring, Summer, Fall)
- The participant answers the following question about him
    1. Does your friend have a pet? (Yes, No, Don't know)
    2. If he/she has a pet, what kind? (Open)

3. If he/she has a pet, what is the name of the pet? (Open)
4. Does your friend have any brother or sister? (Yes, No, Don't know)
5. If your friend has brother(s) and / or sister(s), how many? (Open)
6. If your friend has brother(s) and / or sister(s), what is the name of the elder? (Open)
7. Does your friend practice a sport? (Yes, No, Don't know)
8. If yes, what is this sport? (Open)
9. If which season was your friend born? (Winter, Spring, Summer, Fall, Don't know)

By comparing the answers one gives about his friend and the friend's answers, we are able to compute how many times answers match. This proxies mutual knowledge. In case participants do not know each other, we expect answers to almost never match. Under the null hypothesis that we recruited pairs of participants that were not actual friends, we can assume that they would answer randomly to the questions about their friends. In this case, we further assume that they would give correct answers to open questions ("What is the name of the pet?") with probability 0, and correct answers to closed questions with n possible answers with probability $1/n$. This assumption is conservative: because we let participants simply say that they don't know the answer to open questions, the probability that they give a good answer to such questions is probably lower than $1/n$. The previous assumption implies that, under the null hypothesis, we expect participants to give an average of 1.75 correct answers (S.D.: 5.062) and a maximum of 4 correct answers (with probability 0.031).

On average, participants gave 4.73 (S.D.: 2.33) correct answers about their friends, which is clearly greater than what we would expect under the null hypothesis (t-test, $p < 0.001$). The former suggests that our recruitment procedure was successful in having actual friends coming to the lab.

Measure of the quality and satisfaction of a friendship

We adapt the McGill questionnaire (Mendelson and Aboud, 1999). This test is aimed at measuring positive feelings of responder toward an identified friend. It consists of item about which participant give their opinion on a 9-point likert scale. We selected the following 11 items from the original test:

1. You care a lot about your friend.
2. You like your friend a lot.
3. You want to stay friends for a long time.
4. You prefer your friend over most people you know.
5. You are happy you and your friend are friends.
6. You hope you will stay friends for a long time.
7. You would miss your friend if she/he left.
8. You enjoy him/her as a friend.
9. You are satisfied with your friendship.
10. You think your friendship is a great one.
11. You think your friendship is strong.

Opinion are between strongly disagree (-4) and strongly agree (+4). The scale has an excellent reliability (Cronbach's $\alpha = 0.913$).
In the original study, casual friends had an average score of 2.2 and best friends had an average score of 3.5. In our study, participants had an average score of 2.976 (95% confidence interval : $[2.77; 3.17]$). This suggests that we were successful in recruiting casual friends at least.

A2 Supplementary analyses

Marginal effect of observing misreporting on the decision to play.

Table A2.1 report the outcome of logit regressions for the marginal effect of observing misreport on the decision to play, over experimental condition. The first column presents the simplest specification. In some specifications, we add individual characteristics (gender, risk attitudes, experience with experiment and monthly expenses) and session fixed-effects. The results presented in this table confirm Result 2: observing misreporting reduces the probability that the Audience plays in the treatment, but not in the baseline. Moreover, the marginal effect of observing misreporting is significantly greater in the treatment than in the baseline in every specification.

Table A2.1 – Effect of observing misreporting on the decision to play

	(1) play=1	(2) play=1	(3) play=1	(4) play=1
Baseline	0.169	0.174	0.0953	0.166
	(0.197)	(0.196)	(0.186)	(0.196)
Treatment	-0.314****	-0.324****	-0.304****	-0.360****
	(0.0927)	(0.0905)	(0.0858)	(0.0877)
$p-value$ diff.[a]	0.027	0.021	0.066	0.017
Session F.E.	No	Yes	No	Yes
Ind. char.	No	No	Yes	Yes
N	129	129	129	129
Clusters	40	40	40	40

Robust standard errors clustered at the individual level in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$

Treatment effect on misreporting

Table A2.2 reports the outcome of logit regressions for the treatment effect on misreporting by f_1. The first column presents the simplest, univariate specification. In some specifications, we add individual characteristics (gender, risk attitudes, experience with experiment, and monthly expenses) and session fixed-effects. The results presented

in this table confirms Result 3: there is no significant difference in misreporting by f_1 between condition.

Table A2.2 – Misreporting by f_1s explained by treatment.

	(1)	(2)	(3)	(4)
	f_1 mis.=1	f_1 mis.=1	f_1 mis.=1	f_1 mis.=1
Treatment	-0.0310	-0.299	-0.0465	-0.374
	(0.135)	(0.254)	(0.115)	(0.243)
Session F.E.	No	Yes	No	Yes
Ind. char.	No	No	Yes	Yes
N	129	129	129	129
Clusters	40	40	40	40

Robust standard errors clustered at the individual level in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$

A3 Instructions

(translated from French)

TREATMENT

Thank you for participating in this experiment on decision making. For the length of the session, please shut you phone down, and do not communicate unless we ask you to.

For showing up on time, you already get a 5 € show-up fee. This amount will be added to the payment you will get at the end of the session. During the experiment, you will have to take a number of decisions anonymously. Those decisions can make you earn money. Your earnings will be paid at the end of the session, in private in a separate room.

During the experiment, we will not talk about euros, but about experimental currency units (ECU). The exchange rate is : 1€ = 125 ECU

The experiment is composed of three successive parts. This set of instructions is for the first part only. At the end of each part, you will be given the instructions for the next part.

Should you have any question, please us the call button situated in you cubicle. A experimenter will come to answer to your questions in private.

First Part

The first part is composed of 4 periods.
At the beginning of the first period, the couple of friends who came together at the session will be divided in two. For half of the couple of friends, the program will randomly assign each member of the couple either the role of Friend 1 or the role of Friend 2. A

Friend 1 and a Friend 2 constitute a pair of friends. Each pair of friends is composed of two friends who came together to the session. Each pair remains fixed for the duration of the first part: for each period, Friend 1 and Friend 2 came together to the session.

The second half of the couple of friends will be assigned the role of third party.

For each period, the interactions are between 3 participants: a pair of friends, composed of a friend 1 and a friend 2, and a third party.
For each period, each pair of friends interacts with a different third party. Each pair of friends interacts exactly once with each third party and vice versa.

Before the first period, a playing card is drawn for each pair of friends among Jack of Diamonds, Queen of Diamonds, King of Diamonds, Ace of Diamonds, Joker. The drawing is independent for each pair of friends. The five cards are as likely to be drawn. The drawn card is associated to the pair of friends and this association remains fixed for the length of the first part, that is each pair of friends is associated to the same card for the 4 successive periods. The card associated to the pair is called the “card of the pair". The five cards have exactly the same value.

Friends 1 and Friends 2 are informed, before the beginning of the first period, of their “card of the pair", as shown in the following screen-shot:

Description of a period

At the beginning of each period, the software draws a playing card independently for each pair. This card is called “the winning card". This card can be Jack of Diamonds, Queen of Diamonds, King of Diamonds, Ace of Diamonds, Joker. Each card has the same probability to be drawn. The winning card can be the same from the pair’s card, or it can be different.

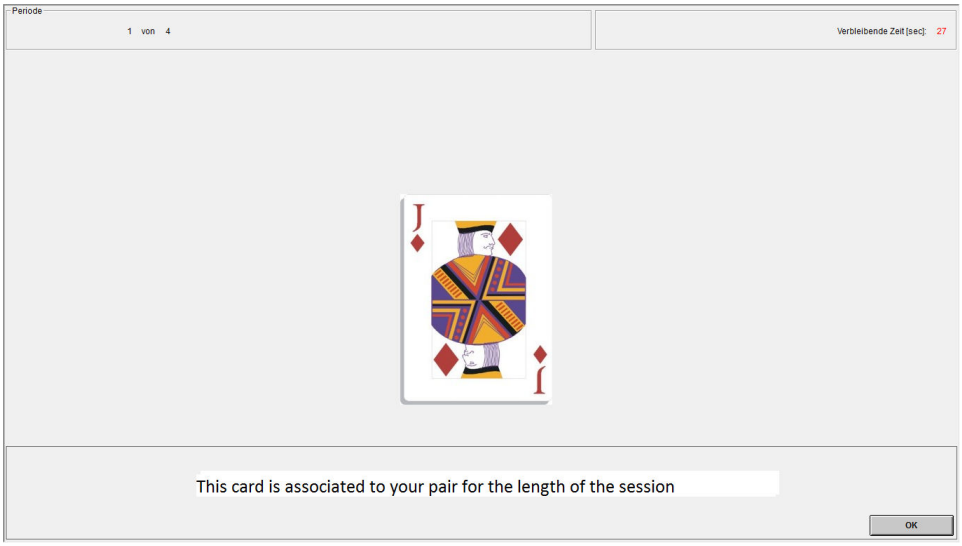


Figure A3.1 – Screenshots for the “card of the pair"

The winning card is displayed on the screen of friends 1 and friends 2 of each pair, as shown in the following screenshot :

After this, friend 1 and friend 2 must report the card of their pair simultaneously and without communicating, as shown in the following screenshot:

Once friend 1 and friend 2 in each pair have declared a card, the third party who is interacting with this pair for this period is informed of:

- The card of the pair.
- The winning card for this period.
- The card reported by friend 1.

The third party then must decide whether he wants to “implement" or not the report of friend 2. However, the third party cannot observe the report of friend 2. Implementing the report of friend 2 means that the report of friend 2 will determine the payoff of the third party and of friend 2 for this period. Not implementing leads to another distribution of payoff, indenpendent of the card of the par and of the winning card, as explained below. The following screenshots represent the decision screen for the third party:

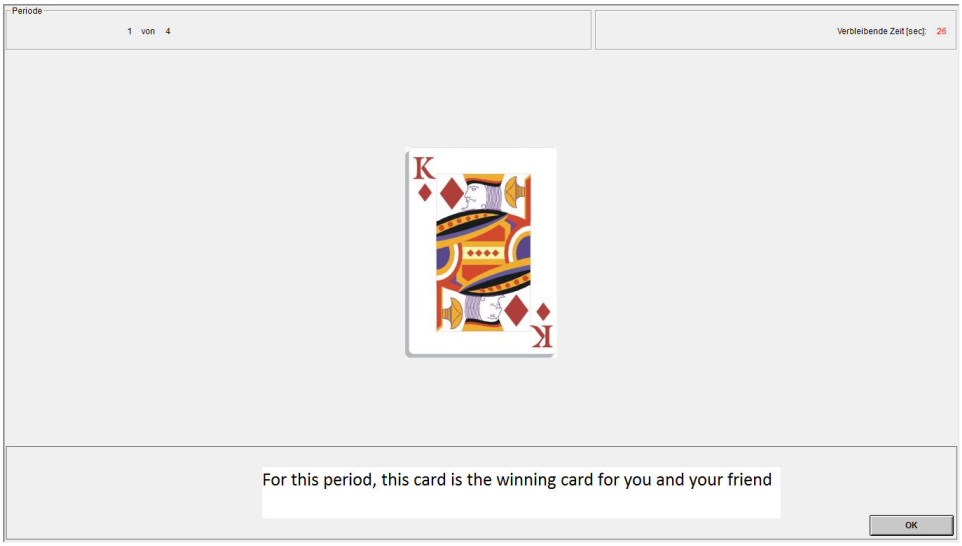


Figure A3.2 – Screenshots for the “winning card"

The decision of the third party ends the period. No further information is revealed. A new period starts automatically. Participants will not be informed of their payoff before the end of the session, if this period is selected for payment.

Determination of payoff

At the end of the session, the software will randomly one of the four period in order to determine the payoff of all the participants in the session for this part. Each period has the same probability to be drawn. Participants are informed of their gains, but of the period selected or of the payoff of other participants.

Friend 1: The payoff of friend 1 is composed of two elements:

- His report
  - If friend 1 declared a card that is not the winning card, he earns 875 UME.
  - If friend 1 declarer the winning card, he earns 1125 UME.
- The decision of the third party:

Figure A3.3 – Screenshots for the “report decision"

- If the third party implements the report of friend 2, friend 1 earns an extra 625 UME.

Friend 2: The payoff of friend 2 depends on the decision of the third party:

- If the third party decides not to implement the report of friend 2, friend 2 earns 750 UME.
- If the third party decides to implement the report of friend 2, the payoff of friend 2 depends on friend 2’s report:
  - If friend 2 reported the winning card, friend 2 earns 1375 UME.
  - If friend 2 reported another card, friend 2 earn 1000 UME.

Third party The payoff of the third party depends on his decision and, if the third party decides to implement, of the report of friend 2:

- If the third party decides not to implement, the third party earns 875 UME.
- If the third party decides to implement, then:

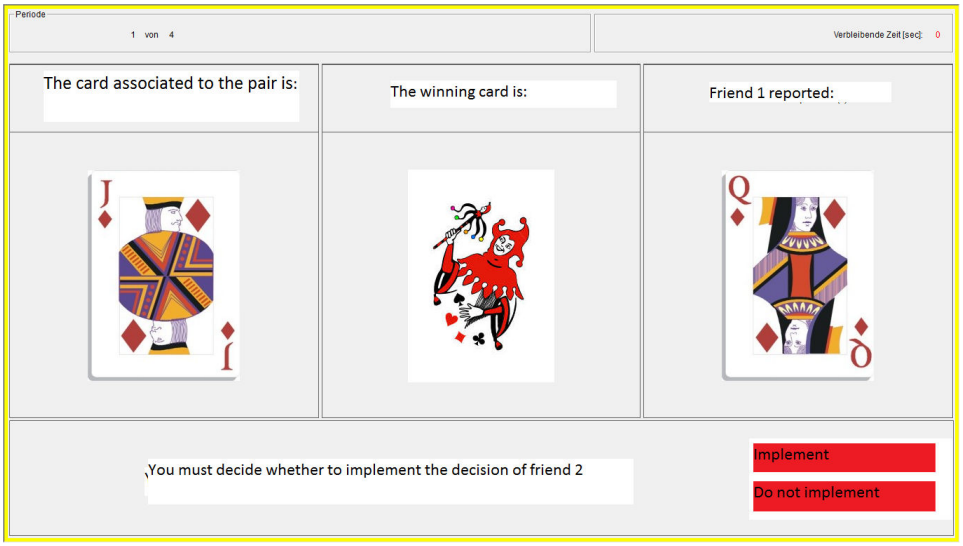


Figure A3.4 – Screenshots: decision screen of the third party"

- if friend 2 reported the card of the pair, the third party earns 1250 UME.
- if fiend 2 reported another card, the third party earns 625 UME.

***

Please read this instructions again. Should you have any question, please use the call button situated in your cubicle. And experimenter will come to answer your questions privately.

***

Part 2

In this part, you must predict the outcome of a coin toss. The coin is balanced: head and tail are equally likely.

In practice:

- You predict the outcome of the toss in your head.
- The software generate the toss and displays the outcome.
- You must indicate whether your prediction was correct.
- If your prediction was correct, you earn an extra 125 UME for this part.
- If your prediction was not correct, you eanr nothing for this part.

Part 3

In this part, some situations from the experiment will be described on your screens. In each of these situations, participants had to choose one of two decisions.

In each case, you must predict the percentage of decisions by participants of the session corresponding to each option.

Below the description of the situation, you will see a slider. You must use this slider to input your predictions.

Payoff for this part

At the end of the session, the software will randomly select one of your predictions. If this prediction is sufficiently close to the actual percentage, you earn money. The following rule is used to compute your earning for this part:

- If your prediction is equal to the true percentage +/- 5 %, you earn 250 ECU.
- If your prediction is equal to the true percentage, +/- 10 %, your earn 187.5 ECU.
- If your prediction is equal to the true percentage, +/- 15 %, your earn 125 ECU.
- If your prediction is equal to the true percentage, +/- 20 %, your earn 62 ECU.
- If your prediction is less accurate, you earn nothing for this part.

Example:

Participants had to choose between option A and option B. 65 % of decisions corresponded to option A (and 35 % to option B).
If you predict:

- 40% of decisions A, you earn nothing,

- 45% of decision A, you earn 62 ECU,
- 50% of decision A, you earn 125 ECU,
- 55% of decision A, you earn 187.5 ECU,
- 60% of decision A, you earn 250 ECU.

***

Please read this instructions again. Should you have any question, please use the call button situated in your cubicle. And experimenter will come to answer your questions privately.

***

BASELINE

Thank you for participating in this experiment on decision making. For the length of the session, please shut you phone down, and do not communicate unless we ask you to.

For showing up on time, you already get a € 5 show-up fee. This amount will be added to the payment you will get at the end of the session. During the experiment, you will have to take a number of decisions anonymously. Those decisions can make you earn money. Your earnings will be paid at the end of the session, in private in a separate room.

During the experiment, we will not talk about euros, but about experimental currency units (ECU). The exchange rate is : € 1 = 125 ECU

The experiment is composed of three successive parts. This set of instructions is for the first part only. At the end of each part, you will be given the instructions for the next part.

Should you have any question, please us the call button situated in you cubicle. A experimenter will come to answer to your questions in private.

First Part

The first part is composed of 4 periods.
At the beginning of the first period, the program will randomly assign participants to roles. There are three possible roles: Person 1, Person 2 or third party.

At the beginning of the first period, the program will randomly form pairs composed of a Person 1 and a Person 2. Each pair remains fixed for the duration of the first part: for each period, each Person 1 forms a pair with the same Person 2 as in the previous

period.

Two friends who came together to the experiment are either both Third Party, or member of two different pairs.

For each period, the interactions are between 3 participants: a pair, composed of a Person 1 and a Person 2, and a third party.
For each period, each pair interacts with a different third party. Each pair interacts exactly once with each third party and vice versa.

Before the first period, a playing card is drawn for each pair among Jack of Diamonds, Queen of Diamonds, King of Diamonds, Ace of Diamonds, Joker. The drawing is independent for each pair. The five cards are as likely to be drawn. The drawn card is associated to the pair and this association remains fixed for the length of the first part, that is each pair is associated to the same card for the 4 successive periods. The card associated to the pair is called the “card of the pair". The five cards have exactly the same value.

Person 1 and Person 2 are informed, before the beginning of the first period, of their “card of the pair", as shown in the following screen-shot:

Description of a period At the beginning of each period, the software draws a playing card independently for each pair. This card is called “the winning card". This card can be Jack of Diamonds, Queen of Diamonds, King of Diamonds, Ace of Diamonds, Joker. Each card has the same probability to be drawn. The winning card can be the same from the pair’s card, or it can be different.

The winning card is displayed on the screen of Person 1 and Person 2 of each pair, as shown in the following screenshot :

After this, Person 1 and Person 2 must report the card of their pair simultaneously and without communicating, as shown in the following screenshot:

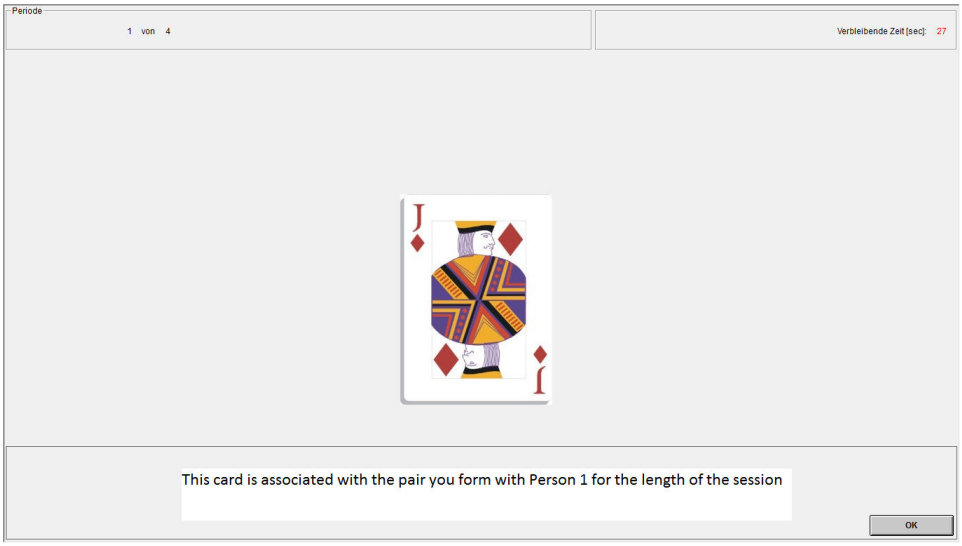


Figure A3.5 – Screenshots for the “card of the pair"

Once Person 1 and Person 2 in each pair have declared a card, the third party who is interacting with this pair for this period is informed of:

- The card of the pair.
- The winning card for this period.
- The card reported by Person 1.

The third party then must decide whether he wants to “implement" or not the report of Person 2. However, the third party cannot observe the report of Person 2. Implementing the report of Person 2 means that the report of Person 2 will determine the payoff of the third party and of Person 2 for this period. Not implementing leads to another distribution of payoff, independent of the card of the pair and of the winning card, as explained below. The following screenshots represent the decision screen for the third party:

The decision of the third party ends the period. No further information is revealed. A new period starts automatically. Participants will not be informed of their payoff before the end of the session, if this period is selected for payment.

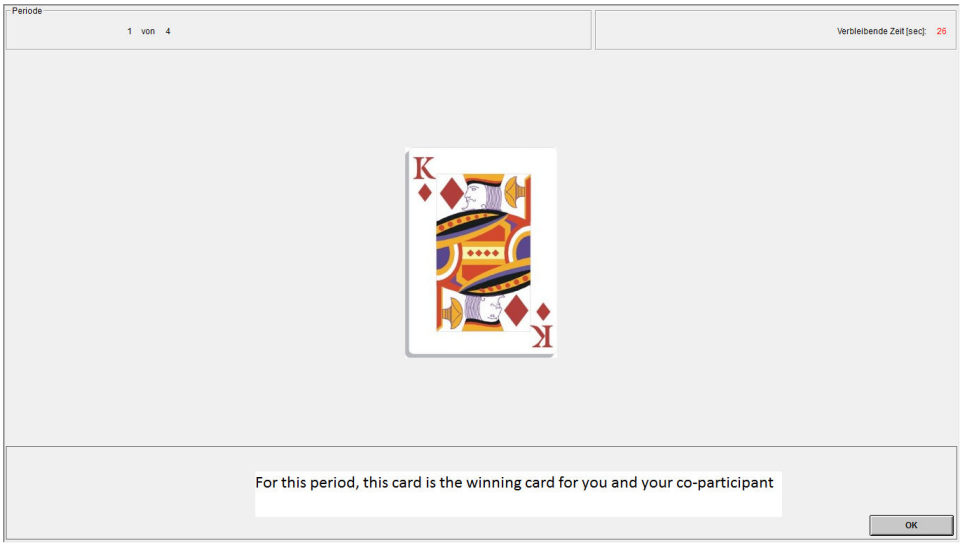


Figure A3.6 – Screenshots for the “winning card"

Determination of payoff At the end of the session, the software will randomly one of the four period in order to determine the payoff of all the participants in the session for this part. Each period has the same probability to be drawn. Participants are informed of their gains, but of the period selected or of the payoff of other participants.

Person 1: The payoff of Person 1 is composed of two elements:

- His report
  - If Person 1 declared a card that is not the winning card, he earns 875 UME.
  - If Person 1 declared the winning card, he earns 1125 UME.
- The decision of the third party:
  - If the third party implements the report of Person 2, Person 1 earns an extra 625 UME.

Person 2: The payoff of Person 2 depends on the decision of the third party:

- If the third party decides not to implement the report of Person 2, Person 2 earns 750 UME.

Figure A3.7 – Screenshots for the "report decision"

- If the third party decides to implement the report of Person 2, the payoff of Person 2 depends on Person 2's report:
  - If Person 2 reported the winning card, Person 2 earns 1375 UME.
  - If Person 2 reported another card, Person 2 earn 1000 UME.

Third party The payoff of the third party depends on his decision and, if the third party decides to implement, of the report of Person 2:

- If the third party decides not to implement, the third party earns 875 UME.
- If the third party decides to implement, then:
  - if Person 2 reported the card of the pair, the third party earns 1250 UME.
  - if Person 2 reported another card, the third party earns 625 UME.

***

Please read this instructions again. Should you have any question, please use the call button situated in your cubicle. And experimenter will come to answer your questions privately.

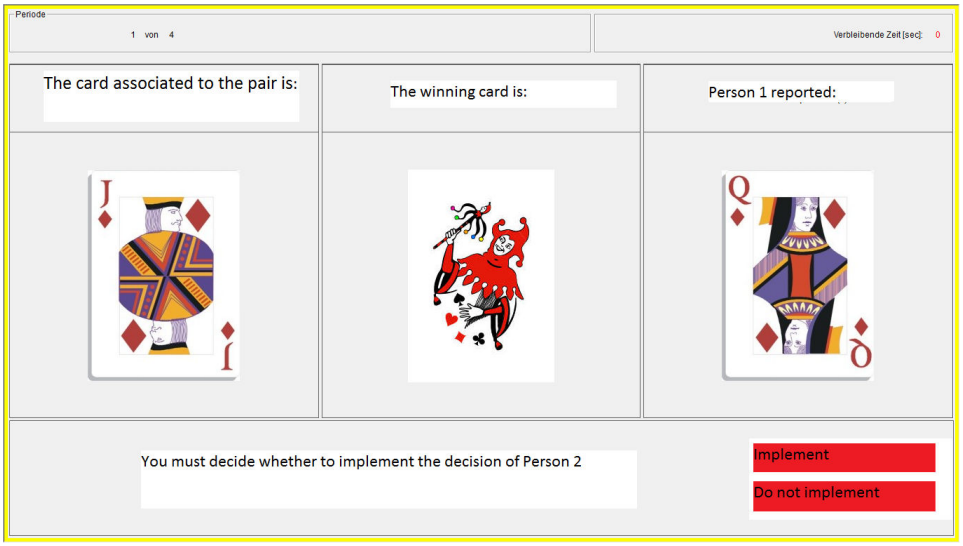


Figure A3.8 – Screenshots: decision screen of the third party"

***

Part 2

In this part, you must predict the outcome of a coin toss. The coin is balanced: head and tail are equally likely.

In practice:

- You predict the outcome of the toss in your head.
- The software generate the toss and displays the outcome.
- You must indicate whether your prediction was correct.
- If your prediction was correct, you earn an extra 125 UME for this part.
- If your prediction was not correct, you eanr nothing for this part.

Part 3

In this part, some situations from the experiment will be described on your screens. In each of these situations, participants had to choose one of two decisions.

In each case, you must predict the percentage of decisions by participants of the session corresponding to each option.

Below the description of the situation, you will see a slider. You must use this slider to input your predictions.

Payoff for this part At the end of the session, the software will randomly select one of your predictions. If this prediction is sufficiently close to the actual percentage, you earn money. The following rule is used to compute your earning for this part:

- If your prediction is equal to the true percentage +/- 5 %, you earn 250 ECU.
- If your prediction is equal to the true percentage, +/- 10 %, your earn 187.5 ECU.
- If your prediction is equal to the true percentage, +/- 15 %, your earn 125 ECU.
- If your prediction is equal to the true percentage, +/- 20 %, your earn 62 ECU.
- If your prediction is less accurate, you earn nothing for this part.

Example:

Participants had to choose between option A and option B. 65 % of decisions corresponded to option A (and 35 % to option B).
If you predict:

- 40% of decisions A, you earn nothing,
- 45% of decision A, you earn 62 ECU,

- 50% of decision A, you earn 125 ECU,
- 55% of decision A, you earn 187.5 ECU,
- 60% of decision A, you earn 250 ECU.

***

Please read this instructions again. Should you have any question, please use the call button situated in your cubicle. And experimenter will come to answer your questions privately.

***

Chapter 3

It does (not) get better: the effect of relative gains and losses on subsequent giving.[1]

1 Introduction

Most people are willing to share their own resources with others, and this contributes to the development of more harmonious societies (Gintis et al., 2003). In the economic literature, this behavior can be motivated by inequality aversion(e.g Fehr and Schmidt, 1999; Bolton and Ockenfels, 2000), the willingness to reward kind intentions (e.g. Charness and Rabin, 2002) or to sustain a good social image (e.g. Andreoni and Bernheim, 2009). These explanations are by nature consequentialist, as they posit that people are willing to share because of the consequences of sharing. Beyond consequences, the context of the decision also conditions the individual willingness to share. This is evidenced by context manipulations in dictator games: for example, adding the possibility to "take" (List, 2007; Bardsley, 2008) or lightly manipulating the legitimacy of dictators (Hoffman et al., 1994; Cherry et al., 2002; Cherry and Shogren, 2008; Carlsson et al., 2013) reduce

[1] Joint work with Julien Benistant.

transfers drastically, whereas minimal moral reminders increase transfers (Brañas Garza, 2007).

In this paper, we identify a new potential determinant of one's willingness to give, that pertains to context rather than to consequences: the salience of counter-factual payoffs, that can be more enviable or less enviable. Our premise is that people will compare their actual payoff to the payoff they could have gotten in a different state of the world: if the payoff compares favourably, they encode it as a gain whereas if it compares unfavourably, they encode it as a loss. This type of situations is commonplace: workers compare the wage raise they get to the one they were expecting (Abeler et al., 2011), people feel worse if they miss an opportunity to increase their payoff than if they were not conscious of this opportunity (Filiz-Ozbay and Ozbay, 2007) and customers are more willing to buy a product when they believe the same product is more expensive elsewhere or if they expected to have to pay a higher price (Heidhues and Kőszegi, 2008). Does getting a payoff lower than what one could have earned reduce one's willingness to give? Symmetrically, does getting a payoff higher than than what one could have earned increase one's willingness to share?

Our laboratory experiment isolates the pure effect of counter-factual potential payoffs on one's willingness to give. In the first part, participants have to perform a real effort task in exchange of a fixed wage. In the second part, they play a dictator game and they have to decide how much of a fixed endowment they are willing to transfer to an other anonymous participant. In a 3×2 between subjects design, we vary the wage for the real effort task (low, medium, high) and the timing of information on the wage (Ref and $NoRef$). In the $NoRef$ conditions, participants are informed upfront of the wage in the real effort task and they are given an envelope containing the corresponding wage before performing the real effort task. In the Ref conditions, they only know the distribution of potential wages and they are informed that the actual wage will be revealed after the real effort task. In all conditions, participants receive an envelope containing the

medium wage from which we take out money in the Ref_{low} condition or add money in the Ref_{high} condition after the revelation of the actual wage. By informing participants of the distribution of the wage, we let them evaluate their actual wage relative to the other possible levels. To isolate the effects of providing information about counter-factual levels of wage, we compare transfers in the Ref conditions to the transfer in the $NoRef$ conditions, for each levels of wage.

As the comparison of the actual wage to the other potential levels may trigger emotional responses that in turn could translate into other-regarding behaviors (e.g Charness and Grosskopf, 2001; Capra, 2004; Kirchsteiger et al., 2006; Andrade and Ariely, 2009; Persson, 2018), we also collected data on the participants' emotional state throughout the experiment, using both declarative measures (Bosman and van Winden, 2002) and physiological measures (Skin Conductance Response). These data are used to assess the role of emotions in our treatment effects.

A natural conceptual framework to model the treatment effects of the experiment are reference-dependent models of preferences (Kahneman and Tversky, 1979; Tversky and Kahneman, 1991; Köszegi and Rabin, 2006) because they assume that individuals perform counter-factual evaluations. Building on this, we develop a simple reference-dependent model of giving in which a participant formulates a reference about his payoff upon being informed of the content of the experiment and then evaluates his actual payoff relative to the reference. The reference payoff is determined by the reference wage and the amount he plans to transfer for this reference wage. The reference wage is assumed to be in-between the high and the low reference wage, so that the high wage is encoded as a gains while the low wage is encoded as a loss. This assumption is consistent with both the status-quo based and the expectation-based formation of the reference point which would be equal to the medium wage, respectively because the envelops contain the medium wage and because the expected wage is equal to the medium wage. Moreover, participants cannot plausibly formulate a reference point

below the low wage or above the high wage. This insures that at least some participants encode the low (resp. high) wage as a loss (gain) while other encode it neither as a loss nor as a gain.[2] The model predicts that participants in the Ref_{low} condition will transfer less than those in the $NoRef_{low}$ because it is a way for them to reduce the gap between their reference payoff and the payoff they actually get. The model also predicts that participants in the Ref_{high} condition will transfer more than those in the $NoRef_{high}$ condition.

The experimental data provide some support to the theory. We find that participants transfer similar amounts in the Ref and $NoRef$ conditions when they earn the medium wage, which suggests that merely being exposed to wage uncertainty has no impact on subsequent transfers. In contrast, loss-averse participants transfer less in the Ref_{low} condition than in the $NoRef_{low}$ while participants transfer more in the Ref_{high} compared to the $NoRef_{high}$ condition, irrespective of whether they are loss-averse. Contrary to the notion that "losses loom larger than gains" (Kahneman and Tversky, 1979), we find no evidence that losses have a greater impact than gains on subsequent giving decisions. As expected, participants in the Ref_{low} condition experience more negative emotions while participants in the Ref_{high} condition experience more positive emotions. Nevertheless, physiological data have very limited explanatory power: the participants who experience higher emotional arousal do not react more strongly to the treatment than those who exhibit little emotional arousal. This suggests a limited role of visceral drives in our experiment.

We contribute to three streams in the literature. First, our results join a large literature showing that other-regarding behavior such as giving are influenced by the context of the decision (e.g. Dana et al., 2007; List, 2007; Bardsley, 2008, are famous examples.). In our experiment, the context of the decision is manipulated by informing participants of the potential payoffs they could have earned in another state of the world. Second, by

[2] Baillon et al. (2018) uncover an important heterogeneity in the way people formulate their reference point. However, they find that the "status-quo" is the modal reference point. Our design is accomodate any reference point that would be between the high and the low wage.

identifying a trade-off between the pro-social preference to give and the preference for avoiding private losses relative to a reference, our results contribute to a more limited literature that studies reference-dependent social preferences. For instance, Charité et al. (2016) provide evidence that dictators are influenced by the reference payoffs of the receivers and Choi (2014) shows that individuals transfer less to charities when their consumption increase less than expected. On the other hand, Buffat (2016) finds that transfers chosen by participant in his lab experiment are not influenced by the level of an exogenous, randomly implemented transfer. This contradicts the predictions of his expectation-based reference dependent model of social preferences. Very closely related to our study, Ockenfels et al. (2015) show that workers who are paid less than their reference wage give less in a subsequent dictator game to the principal who determined their wage. In contrast, in our experiment, recipients cannot be held accountable for the earnings in the real effort task. As a consequence, our results offer a new, complementary perspective on the literature studying how worker perceive their wages. It has been shown than workers compare their wage to the wages of their colleagues(e.g. Card et al., 2012; Bracha et al., 2015) or to the wage they use to have before a cut or a raise (e.g Gneezy and List, 2006; Kube et al., 2012; DellaVigna et al., 2016) and this conditions effort provision and job satisfaction. We identify a new type of reference, that is the potential wages one could have received in a different state of the world. In addition, we show that how one's wage compares to a reference also impacts one's social preferences.

Section 2 describes the experimental design and procedures, Section 3 introduces the theoretical framework and the hypotheses tested in the experiment, Section 4 presents the results and Section 5 concludes.

2 Experimental design and procedures

2.1 Experimental game

Our experimental game consists of three parts. In the first part, participants have to perform a real effort task. In the second part, they play a dictator game. In the third part, we elicit beliefs about the average transfer in the dictator game. Instructions are in Appendix A1.

The real effort task

In the real-effort task, participants have to encrypt a letter into a number 120 times, with the help of a correspondence table displayed on the computer screen. For each new letter, the correspondence table changes, and is randomized at the individual level. If the answer is not correct, they have to enter a new number. Thus, every participants have to solve correctly 120 encryption tasks. We chose this real-effort task because it is both easy to explain and tedious. We fixed the number of letters such that most participants would complete the task in 8 to 10 minutes.

For this task, participants are paid a wage of 5, 10 or 15 euros, depending on the experimental condition. Participants receive an envelop containing either one, two or three €5 banknotes. The envelops were disposed in the cubicles before the session. Our experimental conditions vary the level and timing in which participants are informed of their actual wage for the first part.

The dictator game

In the second part, participants play a dictator game. All participants have to decide how to share a €5 endowment with another anonymous participant. Participants can give any amount between 0 and 5 euros, in increments of 0.1 euro. Participants are informed that they would be randomly matched at the end of the session with another participant. Within each pair, the decision of one participant would be implemented,

while the other participant will be the receiver. This procedure enables us to collect the transfer decisions of every participants in an incentive-compatible way. The transfer decisions in this part constitute our main interest variable.

Belief Elicitation

In the third and final part, participants are asked to estimate the average transfer in their session. They were paid €1 if their prediction was equal to the actual average transfer, plus or minus €0.1.

2.2 Experimental conditions

The experimental conditions vary across participants along two dimensions. Both dimensions are related to the wage for the real effort task. The first dimension varies the level of wage itself. Participants earn €5 in the low conditions, €10 in the medium conditions and €15 in the high conditions.

The second dimension varies the timing in which participants are informed of their actual wage. In the $NoRef$ condition, participants are informed of their actual wage before the real effort task. The envelops placed in the cubicles before the beginning of the session contain their actual and definitive wage. Participants are only informed of the wage they get and that every participants in the session get the same wage, but are not informed that other participants in other sessions get different wages.

In contrast, in the Ref conditions, participants are told at the beginning of the first part that they could earn a wage of either €5 , €10 or €15 with equal probability ($\frac{1}{3}$). They are informed that the actual wage was determined for each sessions before the first session of the experiment, and that they would learn their actual wage at the end of the real effort task. The envelops placed in the cubicles before the beginning of the session contained two €5 banknotes. In the Ref_{low} condition, an experimentalist withdraws one

€5 banknote from the envelops after the actual wage has been revealed to subjects. In the Ref_{high} condition, an experimentalist adds one €5 banknote. In the Ref conditions, the existence of other potential levels of wage is salient. Participants are able to compare the wage they actually get to the two other potential levels, and presumably encode a low wage as a loss and a high wage as a gain. Participants cannot make such counter-factual evaluation in the $NoRef$ conditions.

To measure our treatment effect, we compare transfers in the Ref and $NoRef$ conditions for each level of wage. This allows us to tightly control for potential wealth effects. By comparing transfers in Ref_{low} and in $NoRef_{low}$, we isolate the effect of experiencing a loss. By comparing the transfers in Ref_{high} and in $NoRef_{high}$, we isolate the effect of experiencing a gain. The two medium conditions are necessary to insure that the expected outcome in the Ref conditions is equal to ten. The expectation-based and status-quo based reference wage are thus equal to ten. The medium conditions also allows us to check whether being exposed to uncertainty about the actual wage for the real effort task has an autonomous impact on transfers.[3]

2.3 Additional measures

Measures of emotions In order to explore how the emotional states of participants are affected by our treatment variations, we use two complementary measures.

We record skin conductance, a measure of electro-dermal activity that proxies participants' emotional arousal (Bradley and Lang, 2000). We focus on the electro-dermal response to a specific event: when participants are informed (or reminded in the $NoRef$ conditions) of the actual wage for the real effort task. The response is computed as the average electro-dermal activity in a ten-second window after the revelation (reminder) of the wage from which we subtract an individual baseline computed as the average

[3]Individuals generally dislike being exposed to uncertainty (Loewenstein et al., 2001; Gneezy et al., 2006), and this could impact their behavior in the subsequent dictator game.

activity in a thirty second window at the end of the real-effort task.

Recording electro-dermal response allows us to record any, conscious or unconscious, emotional changes. However, it is uninformative about the valence or the type of emotions experienced by participants. Our second measure of emotions is based on a declarative assessment adapted from Bosman and van Winden (2002) that aimed at associating types and valence to the physiological response. After the dictator game and before the beliefs elicitation, we ask participants to report their emotional state when they learned (or were reminded) of the actual wage for the real-effort task. On ten successive screens, we present to the participants one emotion and they have to report the intensity of this emotion on a 7-point scale, ranging from “no emotion at all" to “high intensity of emotion". The 10 emotions were: anger, surprise, disappointment, joy, elation, jealousy, rage, frustration, irritation and sadness.

Online questionnaire on risk and loss aversion The participants' reaction to falling below or exceeding the reference point might be affected by loss aversion, because participants who want to avoid losses ex-ante might also react more strongly to losses. Therefore, at least one day before the session in the laboratory, participants had to fill out an incentivized online questionnaire aiming at measuring risk and loss aversion according to the method introduced by Eckel and Grossman (2008). Participants had to choose one lottery in two consecutive sets of six binary lotteries displayed on their screen. The expected payoff of the lotteries ranged from €2.8 to €3.6, and variance increased with the expected payoff. The first set is used to elicit risk attitudes in the gain domain, since expected payoffs are always positive. The second set elicits risk attitudes in the loss domains. In order to keep the expected payoff equal, for the second set of lotteries participants were endowed with €4. Participants were told that their choice in one of the two sets of lottery would be chosen randomly for payment, and that they would be informed of the outcome of the lottery at the end of the lab session, when receiving their

payment. We classify participants who choose less risky lotteries in the loss domain than in the gain domain as loss-averse. In addition to this two lottery choices, we collected participants' demographics.

We separate these measures from the core of the experiment to avoid cross-contamination between behavior in the core of the experiment and the choices in the risk and loss aversion elicitation.

2.4 Procedures

The experiment was conducted at the physio-lab of Gate-lab. Participants were recruited using hroot (Bock et al., 2014). The online questionnaire was programmed with LimeSurvey (GmbH., GmbH.) and the laboratory experiment with JAVA. Before the first session, we randomly drew a sequence of conditions thanks to an algorithm that is available upon request. This algorithm ensures that each of the wage levels would be implemented one third of the time. Given this sequence, we planned our sessions by pairs: a Ref and a $NoRef$ condition for the level of wage determined by the algorithm. The use of the algorithm was explained to the participants in the instructions to avoid deception.

In the invitation mailing, participants were informed that we would use physiological measures. We also told them that they would have to fulfill an online questionnaire at least one day before the session. Upon arrival, participants were given an informed consent form detailing the skin conductance measure. In order to participate, they had to explicitly tick the box indicating consent and sign the form. They were informed that they could freely leave the lab if they did not agree, in this case they would receive the amount corresponding to their choice in the online risk aversion task plus five euros. All participants gave their consent.

Table 2.1 – Number of participants by experimental condition

	low wage	medium wage	high wage
$NoRef$	$N = 32$	$N = 30$	$N = 28$
Ref	$N = 34$	$N = 36$	$N = 30$

After consent, an experimentalist took participants to a bathroom where they were invited to wash their hands thoroughly with a special soap. They then drew a sticker from an opaque bag, which indicated the cubicle in which they were invited to seat. An experimentalist installed the electrodes on the index and middle finger of the left-hand of each participant. The temperature was maintained close to 23 degree Celsius in order to improve the accuracy of the electro-dermal measures. An experimentalist then distributed the instructions for the first two parts, these instructions were read aloud by the same experimenter and participants could ask questions.

Overall, 190 participants participated in the experiment. 57 % of the participants are female and the mean age is 23.67. Table 2.1 summarizes the number of participants in each experimental condition.[4]

Sessions were run between June and early October 2018. Our procedure insures that, for any given day in which we planned a session, we would have a $NoRef$ condition for a given wage and the corresponding Ref condition.[5]

Sessions lasted on average 45 minutes. The average payoff is €20.5 (Standard Deviation, SD hereafter: 4.6) including a €5 show-up fee and an average of €3.41 (S.D. 1.9) for the online questionnaire. Table A2.1 in Appendix A2 summarizes the demographic characteristics of participants across conditions.

[4]We invited 12 participants for each sessions, but due to the now-show, we do not have a perfect balance across conditions.

[5]There is one exception: for one session, the software crashed so that we had to trash the data and plan another session some days later to compensate.

3 Theoretical framework and hypotheses

3.1 The model

This section develops a reference-dependent model of giving.

Let us consider an individual i. The payoff of i, π_i, is the sum of two elements: the wage for the real effort task W and of the amount he keeps in the dictator game. This amount is the fraction $0 \leq d \leq 1$ of the endowment E. The profit of the recipient j, π_j, is the sum of the wage for the real effort task W and the amount $(1-d)E$ that the dictator transfers to him. W is determined exogenously and can take three values $W \in \{w_l, w, w_h\}$ with equal probability, with $w_l < w < w_h$ and $w - w_l = w_h - w$. It is common knowledge that i and j receive the same wage.

$$\pi_i = W + dE$$
$$\pi_j = W + (1-d)E$$

i forms references about his payoff ($\overline{\pi_i}$) and the payoff of the receiver ($\overline{\pi_j}$) when he receives the instructions. These references depend on the wage he expects to get in the real effort task, and the hypothetical transfer he would make for this level of wage. In what follows, we assume that w is the reference wage because it corresponds to both the expectation-based and the status-quo based wage. However, the model generates similar predictions with the somewhat weaker assumption that the reference point is between w_l and w_h. We assumes that i's hypothetical transfer corresponds to the transfer i would choose if he would make this choice with no reference in mind, i.e. only with respect to his direct utility function in Equation 3.1.

$$v = m(\pi_i) + \alpha m(\pi_j) \tag{3.1}$$

In this utility function, $m()$ represents' i's valuation of his profit with $m'() > 0$ and $m''() < 0$. $\alpha m()$ represent i's valuation of j's payoff. $\alpha < 1$ captures the fact that i values

his profit more than j's profit. The value of π_i that maximizes i's direct utility function is given by Equation 3.2.

$$m'(\pi_i) = \alpha m'(\pi_j) \tag{3.2}$$

i chooses the transfer d that respects the condition in Equation 3.2. Importantly, a higher wage allows i to increase both his and j's payoff (a more formal proof is presented in Appendix A3). Moreover, Equation 3.2 implies that $\pi_i > \pi_j$ and given that the wage for the real effort part is the same for both participants, participant i transfers less than one half of the endowment in the dictator game.

As explained previously, the hypothetical transfer $\overline{d_w}$ that maximizes the direct utility function for the wage w, determines the references $\overline{\pi_i}$ and $\overline{\pi_j}$.

At the beginning of the dictator game, i is informed of the actual wage for the real effort task, $W \in \{w_l, w, w_h\}$ and chooses his actual transfer d given his utility function. This utility is given by Equation 3.3.

$$\begin{aligned} u = {} & m(\pi_i) + \alpha m(\pi_j) + \mu(m(\pi_i) - m(\overline{\pi_i})) \\ & + \mu(\alpha m(\pi_j) - \alpha m(\overline{\pi_j})) \end{aligned} \tag{3.3}$$

This utility function has the classical features of reference dependent models. $\mu()$ is the gain-loss utility function, which captures the impact of exceeding or falling below the reference level of own and others' profit. An important implication of this utility function is that i evaluates his and j's profit compared to the endogenous references, $\overline{\pi_i}$ and $\overline{\pi_j}$.

Our utility function makes three implicit assumptions common in the literature (see e.g. Köszegi and Rabin, 2006). First, gains and losses are expressed in terms of utility. Second, the gain-loss utility function is additively separable in its arguments. Third, the

gain-loss utility function is the same for both components of the direct utility function. We assume a classical shape for the gain-loss utility function:

$$\mu(X) = \begin{cases} X^{\gamma} & if\ X \geq 0 \\ -\lambda(-X)^{\beta} & if\ X < 0 \end{cases}$$

$\lambda \geq 1$ captures the fact that individuals may be more sensitive to losses than to gains. $0 < \gamma < 1$ and $0 < \beta < 1$ capture diminishing sensitivity respectively in the gain and in the losse domains. Note that the utility function u so defined is continuous and concave in π_i for values of β close to 1, a local extremum of this function is thus a global extremum (proofs are presented in Appendix A4).[6]

Equation 3.4 reports the derivative of the reference-dependent utility function defined in Equation 3.3. The analysis of this derivative at different points allows us to formulate hypotheses on the effect of falling below or exceeding the reference payoff on the subsequent transfer.

$$\begin{aligned} u' = {} & m'(\pi_i) - \alpha m'(\pi_j) \\ & + m'(\pi_i)\mu'(m(\pi_i) - m(\overline{\pi_i})) \\ & - \alpha m'(\pi_j)\mu'(\alpha m(\pi_j) - \alpha m(\overline{\pi_j})) \end{aligned} \tag{3.4}$$

Proposition 1. *If participant i get his reference wage, he transfers the amount he initially planned to transfer. By doing so, he insures that his and j's profits are equal to the references. The intuition is that giving more would lead to a loss in terms of i's own profit which would not be compensated by the gain in j's profit. Symmetrically, giving less would lead to a gain in terms of i's profit that is not sufficient to compensate the loss in terms of j's profit.*

Proof. We study the derivative of the utility function in Equation 3.4 at the point that corresponds to the transfer that maximizes the direct utility function. By definition, at this point $m'(\pi_i) - \alpha m'(\pi_j) = 0$ and $\pi_i = \overline{\pi_i}$ and $\pi_j = \overline{\pi_j}$. Thus, the derivative is null.

[6] Both the bounds on the value of α and β and the fact that β is close to 1 reflects empirical calibrations. See Fox and Poldrack (2014) for an overview of the estimates in the literature.

Since u is concave and continuous, this constitutes the transfers that maximizes i's utility function. □

Proposition 2. *An individual who receives w_l as a wage for the real effort part while having $\overline{w} > w_l$ as a reference transfers less than an individual who receives w_l while having w_l as a reference. The intuition is the following: because i values his own profit more than j's profit, i is more sensitive to the loss of some of his own profit than to the loss of some of j's profit. As a consequence, i reduces his transfers to j in order to reduce the loss in terms of his own profit.*

Proof. Consider an individual who planned to have a final profit $\overline{\pi_i} = \overline{w} + \overline{d_{\overline{w}}}$, but actually receives w_l. We show that this individual transfers less than an individual who planned a profit $\overline{\pi_i} = \overline{w_l} + \overline{d_{\overline{w_l}}}$. To do so, we study the sign of the derivative of the utility function for an expected wage of $\overline{w}$, an actual wage of w_l and a transfer of $\overline{d_{\overline{w_l}}}$. At this point, Equation 3.4 yields:

$$\begin{aligned} u' = {} & m'(\pi_i) - \alpha m'(\pi_j) \\ & + m'(\pi_i)\beta(-\lambda)(-1)(m(\overline{\pi_i}) - m(\pi_i))^{\beta-1} \\ & - m'(\pi_j)\alpha\beta(\lambda)(\alpha m(\overline{\pi_j}) - \alpha m(\pi_j))^{\beta-1} \end{aligned} \tag{3.5}$$

As shown previously, $m'(\pi_i) - \alpha m'(\pi_j) = 0$. The sign of the derivative presented in Equation 3.5 depends only on the sign of $m'(\pi_i)\beta(-\lambda)(-1)(m(\overline{\pi_i}) - m(\pi_i))^{\beta-1} - m'(\pi_j)\alpha\beta(\lambda)(m(\overline{\pi_j}) - m(\pi_j))^{\beta-1})$. Using the fact that $m()$ is concave and that $\pi_i > \pi_j$, we can show that this expression is positive (Some more details are given in Appendix A5). This means that when i expects a given wage but actually receives a lower wage, i transfers less than an individual who expects *and* gets the low wage. Note that the former expression has a direct interpretation. $m'(\pi_i)\beta(-\lambda)(-1)(m(\overline{\pi_i}) - m(\pi_i))^{\beta-1}$ represents the extra utility of keeping slightly more than $\overline{d_{\overline{w_l}}}E$, which closes the gap between the profit i expects and the profit i actually gets. $-(m'(\pi_j)\alpha\beta(\lambda)(m(\overline{\pi_j}) - m(\pi_j))^{\beta-1})$ represents the utility cost of transferring slightly less than $(1 - \overline{d_{\overline{w_l}}})$, which widens the gap between what i expects the recipient would get and the amount the recipient actually gets. Below the reference point, i prefers closing the gap between what he gets and what

he was expecting to get, even if it widens the gap between the recipient's profit and what i expected the recipient would get. □

Proposition 3. *An individual who receives w_h as a wage for the real effort part while having the reference $\overline{W} < w_h$ transfers more than an individual who receives w_h while having the reference w_h. The intuition is that the combination of the concavity of $m()$ and diminishing sensitivity makes the extra, unexpected units of profit more valuable when transferred to the recipient rather than kept.*

Proof. Consider an individual who planned to have a profit $\overline{\pi_i} = \overline{w} + \overline{d_{\overline{w}}}$, but actually receives w_h. This individual transfers more than an individual who planned a profit $\overline{\pi_i} = \overline{w_h} + \overline{d_{\overline{w_h}}}$. To do so, we study the sign of the derivative of the utility function for an expected wage of $\overline{w}$, an actual wage of w_h and a transfer of $\overline{d_{\overline{w_h}}}$. At this point, the Equation 3.4 yields:

$$\begin{aligned} u' = {} & m'(\pi_i) - \alpha m'(\pi_j) \\ & + m'(\pi_i)\gamma(m(\pi_i) - m(\overline{\pi_i}))^{\gamma-1} \\ & - m'(\pi_j)\alpha\gamma(\alpha m(\pi_j) - \alpha m(\overline{\pi_j}))^{\gamma-1} \end{aligned} \tag{3.6}$$

The sign of the derivative depicted in Equation 3.6 is determined by the sign of $m'(\pi_i)\gamma(m(\pi_i) - m(\overline{\pi_i}))^{\gamma-1} - m'(\pi_j)\alpha\gamma(\alpha m(\pi_j) - \alpha m(\overline{\pi_j}))^{\gamma-1}$. Using a similar argument than previously, we can show that this expression is negative. This means that when i expects a given wage but actually receives a higher wage, i transfers more than an individual who expects *and* gets the high wage. The intuition for this result is that, because $m()$ is concave, and because sensitivity to gains diminishes with gains, and that i's profit is always greater than j's, a gain in terms of π_j procures a greater utility to i than a gain in terms of his own payoff. □

3.2 Hypotheses

The previous theoretical framework leads to the following hypotheses:

Hypothesis 1: Participants transfer the same amount in the Ref_{med} and $NoRef_{med}$ condition. This follows from Proposition 1.

Hypothesis 2: Participants in the Ref_{low} condition transfer less than participants in the $NoRef_{low}$. This follows from proposition 2.

Hypothesis 3: Participants in the Ref_{high} condition transfer more than participants in the $NoRef_{high}$ condition. This follows from proposition 3.

4 Results

4.1 Behavioral results

The average transfer in the dictator game is €1.37 (S.D. 1.40). This represents 27% of the €5 endowment, which is well in line with the literature.[7] Table 4.2 reports the average transfer of participants across conditions.

Table 4.2 – Average transfers by experimental condition

	Wage=5	Wage=10	Wage=15
$NoRef$	1.64 (N=32, S.D.= 1.52)	1.14 (N=30, S.D=1.14)	1.32 (N=28, S.D.=1.37)
Ref	1.28 (N=34, S.D.=1.53)	1.18 (N=36, S.D.=1.51)	1.69 (N=30, S.D.=1.28)
p-value	0.26	0.94	0.29
Overall	1.45 (N=66, S.D.=1.52)	1.16 (N=66, S.D.=1.34)	1.51 (N=58, S.D.=1.32)

Notes: N is the number of observations, S.D. is the standard deviation. $p-values$ for two-sided Mann-Whitney tests.

While the tendencies match our predictions, none of the comparisons between the $NoRef$ and Ref conditions reach statistical significance.

[7] In a meta-analysis including more than 130 papers and 600 treatments, Engel (2011) finds that dictators give on average 28% of their endowment.

We ran regressions explaining transfers, pooling all the data from the experiment. We interact a dummy variable indicating the Ref conditions, with a categorical variable indicating the wage level. In Table 4.3, we report the marginal treatment effect for the different levels of wage. Models (1) and (2) are left-censored Tobit models (38% of the transfers are equal to zero). Models (3) and (4) are Logit models, explaining the decision to transfer zero. All models include a categorical variable accounting for month fixed effect.[8] Models (2) and (4) include individual characteristics. We control for gender, age, the number of previous participations in economic experiments, risk aversion, and include dummies indicating whether the participant is loss averse[9] and whether the participant is a student. Robust standard errors are clustered at the session level.[10]

The outcome of these regressions supports the three following behavioral results:

Result 1: For a €10 wage, the Ref treatment has no impact on transfers. This result supports Hypothesis 1. None of the parameters capturing treatment effect for the €10 wage is significant in the regressions. This result suggests that being merely exposed to payoff uncertainty does not impact the participants' propensity to share.[11] As a consequence, the treatment effects potentially found for the two other levels of wage are not confounded by the effect of being exposed to uncertainty.

Result 2: For a €5 wage, the "Ref" treatment reduces transfers of loss-averse participants only. The data provide mixed-support to hypothesis 2. The treatment effect is

[8] We ran sessions in early summer which corresponds to the end of the academic year, and in early fall which corresponds to the beginning of the academic year. The participants likely differ both on observable and non-observable characteristics, which is why we control for month fixed-effects.

[9] Participants who chose a lottery with less variance in the loss domain than in the gain domain in the online questionnaire are categorized as loss averse.

[10] There are several reasons to cluster standard errors at the session level even in one-shot experiments (see e.g. Fréchette, 2012). For instance, the experimental procedure to set-up the physiological measures implied interactions between participants and with the experimenter that are hardly measurable and made the room composition in terms of gender salient which can impact behavior (Castillo et al., 2015).

[11] One could have expected lower transfers in the Ref condition than in the $NoRef$ condition for the €10 wage because individuals dislike being merely exposed to uncertainty (see the "uncertainty effect" presented in Gneezy et al., 2006).

Table 4.3 – Treatment effects over the different levels of wage.

	(1) Transfer	(2) Transfer	(3) Transfer=0	(4) Transfer=0
Treatment effect at:				
Wage=5	-0.558*	-0.379	0.165**	0.146
	(0.319)	(0.410)	(0.078)	(0.093)
Wage=10	0.009	-0.175	0.028	0.080
	(0.286)	(0.292)	(0.069)	(0.067)
Wage=15	0.614***	0.816****	-0.163***	-0.199***
	(0.230)	(0.208)	(0.050)	(0.069)
Month F.E.	YES	YES	YES	YES
Individual Char.	NO	YES	NO	YES
N	190	190	190	190

Robust standard errors clustered at the session level in parentheses. (1) and (2) are Tobit models. (3) and (4) are Logit models. Marginal effects reported. Treatment effect corresponds to the marginal effect of the *Ref* condition. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$

always in the expected direction, but reaches statistical significance only in the models (1) and (3) that do not control for individual characteristics.

As €5 is the lowest potential wage, it is probably encoded as a loss by participants. Loss-averse participants transferred less than non loss-averse participants in the $NoRef_{low}$ condition (1.7 *vs.* 0.69, MW: $p = 0.056$), which suggests that loss-averse participants are more sensitive to the treatment. To assess this claim, we run regressions similar to those reported in Table 4.3, except that we add a third term to the interactions between treatment and the level of the wage: a dummy variable indicating that the participant is loss averse. The estimates are reported in Table 4.4.

We conclude that the effect of experiencing a loss on the transfer decision is more marked for loss-averse participants. This result has an intuitive interpretation: those who try to avoid losses (as evidenced by their choices in the online questionnaire) are also those who want to “chase" losses once they experience them. Here, the only way in

Table 4.4 – Treatment effect on loss averse *vs.* not loss averse participants, at wage=€5.

	(1) Transfer	(2) Transfer	(3) Transfer=0	(4) Transfer=0
Treatment effect, at wage=5 for:				
Not loss averse participants	0.013	0.125	0.069*	0.048
	(0.240)	(0.384)	(0.036)	(0.056)
Loss averse participants	-1.185***	-0.983***	0.237**	0.232
	(0.452)	(0.299)	(0.120)	(0.156)
$p-value$ diff.	0.011	<0.001	0.094	0.21
Month F.E.	YES	YES	YES	YES
Individual Char.	NO	YES	NO	YES
N	190	190	190	190

Robust standard errors clustered at the session level in parentheses. (1) and (2) are Tobit models. (3) and (4) are Logit models. Marginal effects reported. Treatment effect corresponds to the marginal effect of the Ref condition. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$

which they can chase losses in terms of their own profit is by transferring less money as dictators.

Result 3: For the €15 wage, participants transfer larger amounts in the Ref condition compared to the $NoRef$ condition. This result supports Hypothesis 3. The treatment effects for a wage of €15 are large and significant in every specification.

4.2 Emotions

We now turn to the analysis of the emotion data. Experiencing losses may trigger negative emotions that could in turn reduce transfers, while experiencing gains may trigger positive emotions that could in turn increase transfers. In what follows, we asses the role of emotions in the treatment effects exposed above.

We first analyze the physiological data.[12] We focus on the electro-dermal activity consecutive to the revelation (or reminder in the $NoRef$ conditions) of the actual wage

[12]We drop 22 individual observations from the analysis because of the poor quality of their electro-dermal signal, which represents 11.5% of our sample.

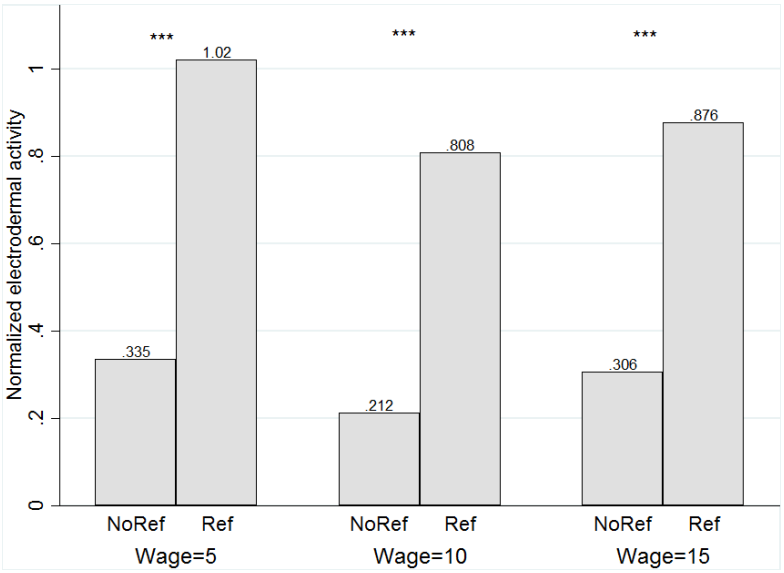


Figure 4.1 – Electrodermal activity across conditions
Notes: *** indicates $p < 0.01$ in a Mann-Whitney two-sided test.

for the real effort task. Figure 4.1 plots the normalized activity.

Electro-dermal activity is always greater in the Ref than in the $NoRef$ condition (MW: $p - values$ are lower than 0.001 for every levels of wage). This shows that the revelation of the wage generated physiological response. However, electro-dermal activity does not differ across level of wage, nor for the Ref conditions (KW: $p = 0.713$), neither for the $NoRef$ conditions (KW: $p = 0.749$).

In order to associate a valence to the emotional arousal identified in the previous paragraph, we now turn to the analysis of self-reported data. We separate the item of the questionnaire in “positive emotions" (Joy, elation, surprise) and “negative emotions" (Anger, disappointment, rage, irritation, jealousy, frustration, sadness). For each individuals, we average the responses to “positive emotions" items and to “negative emotions" items, which gives an individual index of “positive emotions" and an individual index

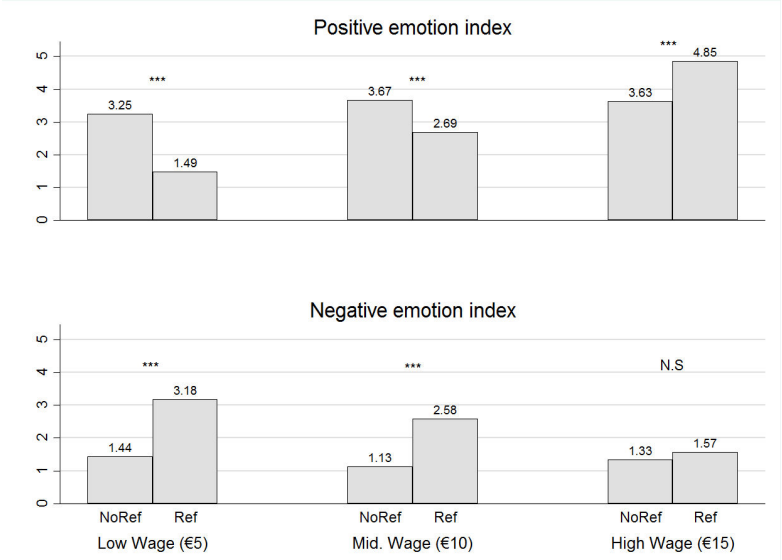


Figure 4.2 – Indexes of positive (top) and negative emotions across conditions
Notes: Index of positive emotions: individual average self-report on a 1-7 scale for Joy, Elation and Surprise. Index of negative emotions: individual average self-report on a 1-7 scale for anger, disappointment, rage, irritation, jealousy, frustration and sadness. *** indicates $p < 0.01$ in a Mann-Whitney two-sided test.)

of "negative emotions".[13] Figure 4.2 plots the positive (top panel) and negative (bottom panel) emotion indexes.

In the $NoRef$ conditions, the wage level does not impact positive emotions (Kruskal-Wallis test, KW hereafter: $p = 0.327$) and mildly impacts negative ones (KW: $p = 0.08$). In the Ref conditions, both positive and negative emotions strongly depend on the level of wage (KW: $p < 0.001$ for both positive emotions, and negative emotions).

Both positive and negative self-reported emotions differ between the $NoRef$ and the Ref conditions. Participants express more negative emotions and less positive emotions when they know that they could have earned more than the wage they get (two first comparisons of the top and bottom panels). In contrast, when they get the maximum

[13]The questionnaire has a good reliability overall (Cronbach's alpha=0.84). In Appendix A6, more details are given about the structure of correlation between the items from the questionnaire.

wage, participants express more positive emotions when they are informed of the other potential wages than when they are not. In both the Ref_{low} and Ref_{mid}, the highest reports were for disappointment, and disappointment was greater in the Ref_{low} than in the Ref_{high} (5.14 *vs.* 3.58, MW: $p < 0.001$). In the other conditions, the highest reports were always for joy, and the highest joy was reported in the Ref_{high} condition (5.32). Figure A6.1 in A6 details the average reports for each emotions, across all conditions.

By combining the data from the self-reported and physiological measures of emotion, we can conclude that the salience of the counter-factual wages influenced the emotional state of participants: (i) when participants earn the low wage, informing them of the other wages triggers negative emotions, in particular disappointment ; (ii) when participants earn the high wage, informing them of the other potential wages triggers positive emotions, in particular joy.

In order to explore the role of the emotions experienced at the time of revelation in the treatment effects, we test whether emotional arousal correlates with the decisions in the trust game in the Ref conditions. We add a third term to the interactions between treatment and the level of wage used in Table 4.3: the normalized physiological measure of emotion. Table 4.5 reports the marginal effects of the normalized physiological measure of emotion for the different levels of wage conditional on being in the treatment. In models (1) and (2), we estimate Tobit models. In models (3) and (4) we estimate Logit models. Models (1) and (3) include month fixed effects. Model (2) and (4) also control for the individual characteristics used in Table 4.3.

We find no evidence that emotional arousal impacts transfers. This suggests that the emotions triggered by the treatment do not translate into behavior.

Table 4.5 – Marginal effects of physiological arousal on transfers over the different level of wage.

	(1) Transfer	(2) Transfer	(3) Transfer=0	(4) Transfer=0
effect of electro-dermal act. at:				
Wage=5	-0.418	-0.227	0.015	-0.017
	(0.914)	(0.840)	(0.169)	(0.147)
Wage=10	0.260	0.081	-0.085	-0.044
	(0.928)	(0.927)	(0.179)	(0.195)
Wage=15	-0.233	-0.185	0.005	-0.004
	(0.488)	(0.414)	(0.093)	(0.086)
Month F.E.	YES	YES	YES	YES
Individual Char.	NO	YES	NO	YES
N	168	168	168	168

Robust standard errors clustered at the session level in parentheses. (1) and (2) are Tobit models. (3) and (4) are Logit models. Marginal effects reported. Treatment effect corresponds to the marginal effect of the *Ref* condition. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, **** $p < 0.001$

5 Discussion and conclusion

In this paper, we tested whether the salience of better or worse counter-factual wages in a real effort task affects giving in a subsequent dictator game. We found that making salient that the wage could have been lower increases the transfer of participants who receive the high wage. Symmetrically, participants who receive the low wage transfer less when they are informed of the other potential levels of wage than when they are not, but the effect is concentrated on loss-averse participants. Our results do not support the intuition that losses have greater consequences than gains.

The effects are found by varying the information about the counter-factual levels of wage between participants, holding the actual wage constant. As a consequence, wealth effects do not confound our results. By design, we also mute reciprocity concerns since the recipients in the dictator game cannot be held accountable for the actual wage in the real effort task. In addition, wage uncertainty does not drive our results, since the transfers of participants who get the reference wage do not vary with the timing in which

they are informed of their actual wage. The physiological and self-reported emotional data show that our treatments triggers emotional reactions, but these reactions have little explanatory power. A higher emotional arousal at the time of wage revelation (or reminder) does not affect behavior in the dictator game. This shows that our results are not merely an avatar of the effect of emotions or mood on other-regarding behaviors (On this topics, see e.g Charness and Grosskopf, 2001; Capra, 2004; Kirchsteiger et al., 2006; Pérez-Dueñas et al., 2018). A possible interpretation of our results is that earning a wage that compares favourably to other possible levels of wage is encoded as a gain by participants and dictators tend to share gains. On the other hand, earning a wage that compares unfavorably with other possible levels is encoded as a loss and dictators reduce their transfers in order to compensate for this loss.

From a methodological stand point, some features of our experiment call for discussion. First, narrow bracketing (e.g. Read et al., 1999; Thaler, 1999) might have attenuated the effects because we gave an extra endowment for the dictator game instead of having the gains or losses affect the endowment directly. Possibly, some participants allocate the gain in the real effort task and the gain in the dictator game to two separate mental accounts and separate the decisions in these two parts. Second, participants in our experiment seem to perform counter-factual evaluation of their wage. The results are consistent with the hypothesis that participants compare their wage to the medium wage, which is both salient and the expectation-based and the status-quo-based reference wage. However, our aggregate results might hide individual heterogeneity in how the counter-factual evaluation is performed (see e.g. Baillon et al., 2018). As an illustration, it is possible that some participants encode the medium wage as a gain because their reference is the low wage, while some may encode it as a loss because their reference is the high wage. A possible extension of our research is to vary the number of potential wages within participants in order to refine our understanding of counter-factual evaluation.

Our results contribute to the literature studying social preferences, and more specifically how the context of the decision impacts giving (e.g. List, 2007; Bardsley, 2008). Our manipulation allows to observe the transfers of dictators whose payoffs differ from their reference, and the data point to the existence of a trade-off between social preferences and the natural tendency of loss averse people to chase losses: losses relative to a reference payoff reduce the concern for other while gains increase it. Self-commitment to act altruistically seems to yield in front of the disutility linked to incurring private losses.

Our results also speaks to the literature studying how workers perceive their wages relative to some references. Workers are influenced by how their wage compares to their peers' wages, with consequences for effort provision (Bracha et al., 2015) and job satisfaction (Card et al., 2012). Workers also compare their actual wage to the wage they use to have, as evidenced by the literature on the effects of wage cuts and raises on effort provision (e.g Gneezy and List, 2006; Kube et al., 2012; DellaVigna et al., 2016). The given interpretation is that workers reciprocate to the principal's intentions by adjusting their level of effort. In our experiment, we mute inequality concerns, as all the participants in a given session earn the same wage ; and reciprocity, as the recipient is not accountable for the actual level of wage. As a consequence, we provide evidence that workers may compare their wage to the wages they could have received in a different state of the world, independently of social comparison or wage dynamic. Moreover, losses and gains in terms of wage affect social preferences, rather than effort provision. This means that how workers evaluate their wage relative to a reference might impact the social capital of firms, and promote or impede cooperation between workers.

Bibliography

Abeler, J., A. Falk, L. Goette, and D. Huffman (2011). Reference Points and Effort Provision. *American Economic Review 101*(April), 470–492.

Andrade, E. B. and D. Ariely (2009). The enduring impact of transient emotions on decision making. *Organizational Behavior and Human Decision Processes 109*(1), 1–8.

Andreoni, J. and B. D. Bernheim (2009). Social image and the 50-50 norm: A theoretical and experimental analysis of audience effects. *Econometrica 77*(5), 1607–1636.

Baillon, A., H. Bleichrodt, and V. Spin (2018). Searching for the reference point. *Management Science*.

Bardsley, N. (2008, Jun). Dictator game giving: altruism or artefact? *Experimental Economics 11*(2), 122–133.

Bock, O., I. Baetge, and A. Nicklisch (2014). hroot: Hamburg Registration and Organization Online Tool. *European Economic Review 71*(C), 117–120.

Bolton, G. E. and A. Ockenfels (2000). Erc: A theory of equity, reciprocity, and competition. *The American Economic Review 90*(1), 166–193.

Bosman, R. and F. van Winden (2002). Emotional Hazard in a Power-to-Take Experiment. *The Economic Journal 112*(476), 147–169.

Brañas Garza, P. (2007). Promoting helping behavior with framing in dictator games. *Journal of Economic Psychology 28*(4), 477 – 486.

Bracha, A., U. Gneezy, and G. Loewenstein (2015). Relative pay and labor supply. *Journal of labor economics 33*(2), 297–315.

Bradley, M. M. and P. J. Lang (2000). Measuring emotion: Behavior, feeling, and physiology. In *Cognitive neuroscience of emotion.*, Series in affective science., pp. 242–276. New York, NY, US: Oxford University Press.

Buffat, J. (2016). Charitable preferences and expectation-based reference dependence: An experiment on charitable giving.

Capra, C. M. (2004). Mood-driven behavior in strategic interactions. *The American Economic Review 94*(2), 367–372.

Card, D., A. Mas, E. Moretti, and E. Saez (2012). Inequality at work: The effect of peer salaries on job satisfaction. *American Economic Review 102*(6), 2981–3003.

Carlsson, F., H. He, and P. Martinsson (2013, Jun). Easy come, easy go. *Experimental Economics 16*(2), 190–207.

Castillo, M., G. Leo, and R. Petrie (2015). Room composition effects on risk taking by gender.

Charité, J., R. Fisman, and I. Kuziemko (2016). Reference points and redistributive preferences: Experimental evidence. *NBER Working Paper No. 21009*.

Charness, G. and B. Grosskopf (2001). Relative payoffs and happiness: an experimental study. *Journal of Economic Behavior & Organization 45*(3), 301–328.

Charness, G. and M. Rabin (2002). Understanding social preferences with simple tests. *The Quarterly Journal of Economics 117*(3), 817–869.

Cherry, T. L., P. Frykblom, and J. F. Shogren (2002). Hardnose the dictator. *The American Economic Review 92*(4), 1218–1221.

Cherry, T. L. and J. F. Shogren (2008). Self-interest, sympathy and the origin of endowments. *Economics Letters 101*(1), 69 – 72.

Choi, S. (2014). Reference-dependent preferences and charitable giving.

Dana, J., R. A. Weber, and J. X. Kuang (2007, Oct). Exploiting moral wiggle room: experiments demonstrating an illusory preference for fairness. *Economic Theory 33*(1), 67–80.

DellaVigna, S., J. A. List, U. Malmendier, and G. Rao (2016). Estimating social preferences and gift exchange at work. Technical report, National Bureau of Economic Research.

Eckel, C. C. and P. J. Grossman (2008). Forecasting risk attitudes: An experimental study using actual and forecast gamble choices. *Journal of Economic Behavior & Organization 68*(1), 1 – 17.

Engel, C. (2011). Dictator games: A meta study. *Experimental Economics 14*(4), 583–610.

Fehr, E. and K. M. Schmidt (1999). A theory of fairness, competition, and cooperation. *The quarterly journal of economics 114*(3), 817–868.

Filiz-Ozbay, E. and E. Y. Ozbay (2007). Auctions with anticipated regret: Theory and experiment. *American Economic Review 97*(4), 1407–1418.

Fox, C. R. and R. A. Poldrack (2014). Appendix: Prospect theory and the brain. In P. W. Glimcher and E. Fehr (Eds.), *Neuroeconomics (Second Edition)* (Second Edition ed.)., pp. 533 – 567. San Diego: Academic Press.

Fréchette, G. R. (2012). Session-effects in the laboratory. *Experimental Economics 15*(3), 485–498.

Gintis, H., S. Bowles, R. Boyd, and E. Fehr (2003). Explaining altruistic behavior in humans. *Evolution and human Behavior 24*(3), 153–172.

GmbH., L. Limesurvey: An open source survey tool. LimeSurvey GmbH.

Gneezy, U. and J. A. List (2006). Putting behavioral economics to work: Testing for gift exchange in labor markets using field experiments. *Econometrica 74*(5), 1365–1384.

Gneezy, U., J. A. List, and G. Wu (2006). The uncertainty effect: When a risky prospect is valued less than its worst possible outcome. *The Quarterly Journal of Economics 121*(4), 1283–1309.

Heidhues, P. and B. Kőszegi (2008). Competition and price variation when consumers are loss averse. *American Economic Review 98*(4), 1245–68.

Hoffman, E., K. McCabe, K. Shachat, and V. Smith (1994). Preferences, property rights, and anonymity in bargaining games. *Games and Economic Behavior 7*(3), 346 – 380.

Kahneman, D. and A. Tversky (1979). Prospect theory: An analysis of decision under risk. *Econometrica 47*(2), 263–291.

Kirchsteiger, G., L. Rigotti, and A. Rustichini (2006). Your morals might be your moods. *Journal of Economic Behavior & Organization 59*(2), 155–172.

Köszegi, B. and M. Rabin (2006). A model of reference-dependent preferences. *The Quarterly Journal of Economics 121*(4), 1133–1165.

Kube, S., M. A. Maréchal, and C. Puppe (2012). Do wage cuts damage work morale? evidence from a natural field experiment. *Journal of the European Economic Association 11*(4), 853–870.

List, J. A. (2007). On the Interpretation of Giving in Dictator Games. *Journal of Political Economy 115*(3), 482–493.

Loewenstein, G. F., E. U. Weber, C. K. Hsee, and N. Welch (2001). Risk as feelings. *Psychological bulletin 127*(2), 267.

Ockenfels, A., D. Sliwka, and P. Werner (2015). Bonus payments and reference point violations. *Management Science 61*(7), 1496–1513.

Pérez-Dueñas, C., M. F. Rivas, O. A. Oyediran, and F. Garcìa-Torres (2018). Induced negative mood increases dictator game giving. *Frontiers in Psychology 9*, 1542.

Persson, E. (2018). Testing the impact of frustration and anger when responsibility is low. *Journal of Economic Behavior & Organization 145*, 435 – 448.

Read, D., G. Loewenstein, and M. Rabin (1999). Choice bracketing. *Journal of Risk and Uncertainty 19*(1/3), 171–197.

Thaler, R. H. (1999). Mental accounting matters. *Journal of Behavioral decision making 12*(3), 183–206.

Tversky, A. and D. Kahneman (1991). Loss aversion in riskless choice: A reference-dependent model. *The Quarterly Journal of Economics 106*(4), 1039–1061.

A1 Instructions

Thank you for participation in this experiment on decision making. You are not allowed to communicate with other participants for the length of the experimental session. Your decisions are anonymous. In addition, because we record your skin conductance, please avoid making noise throughout of the session.

The session is composed of three parts. The instructions hereafter describe the content of Parts 1 and 2. The instructions for the third part will be displayed on your screen.

Your decisions can make you earn money. Your payoffs, except for Part 1, will be payed in cash, in private at the end of the session and in a separate room.

We will read the instructions together. Should you have any question, raise your hand. An experimentalist will come to you.

Part 1

For this part, you have to perform 120 encoding tasks. The task consists in encoding a letter from the alphabet into a number. To know the number corresponding to the letter, you must use a correspondence table that is displayed on your screen. After inputting the letter in the dedicated field, you must validate your answer. If your answer is correct, you go on to the next letter; otherwise you have to input a new number. For each letter, the correspondence table changes.

[The next paragraph depends on the conditions.]

[NoRef condition (wage=€5).]

For this task, you will receive a payoff. An envelop containing €5 was placed on your desk for this purpose. The content of the envelop of every participants in the session is the same. The content of the envelop is your payoff for the first part. The content of the envelop will be reminded to you at the end of Part 1.

***

[NoRef condition (wage=€10).]

For this task, you will receive a payoff. An envelop containing €10 was placed on your desk for this purpose. The content of the envelop of every participants in the session is the same. The content of the envelop is your payoff for the first part. The content of the envelop will be reminded to you at the end of Part 1.

***

[NoRef condition (wage=€15).]

For this task, you will receive a payoff. An envelop containing €15 was placed on your desk for this purpose. The content of the envelop of every participants in the session is the same. The content of the envelop is your payoff for the first part. The content of the envelop will be reminded to you at the end of Part 1.

***

[Ref conditions.]

For this task, you will receive a payoff. An envelop containing €10 was placed on your desk for this purpose. Nevertheless, at the end of the encoding tasks, a random draw will determine your definitive payoff for this part. This random draw was done at

the beginning of the set up of the experiment, before the first session, for every sessions. There is 1 chance out of 3 that the content of the envelop remain unchanged, 1 chance out of 3 that we withdraw €5 from the envelop and 1 chance out of 3 that we add €5 in your envelop.

The random draw is at the session level, so that the content of every participant in a given session is the same. The random draw does not depend on your behavior or on the behavior of other participants. The final content of the envelop is your payoff for Part 1.

***

Part 2

In this part, the decisions involve a person A and a person B. Person A receives a €5 endowment. She decides how much she wants to transfer to Person B. Every amount between 0 and 5 are possible, rounded to 10 cents. Person A keeps the amount she does not transfer.

Person B does not receive an endowment. She earns the amount that Person A transferred. She has no decision to make.

Every participants in the session take the decision as Person A. At the end of the session, the software will randomly match participants of the session in pairs. For each pair, the software will randomly select a Person A and a Person B. For Person A, her decision will determine her payoff for this part. For Person B, the decision of the Person A to whom she is matched will determine her payoff for this part.

Payoff for Part 2 will be payed at the end of the session.

A2 Demographic characteristics

Table A2.1 – Demographic characteristics.

Sessions	Overall	Summer	Fall	$p-value$
Prop. Female	0.57	0.56	0.58	0.878
Prop. Student	0.82	0.66	0.91	< 0.001
Age	23.67	26.9	21.95	< 0.001
Participation	3.48	6.43	1.91	< 0.001
N	190	66	124	-
N Session	18	6	12	-
Wage=5	6	$NoRef$:2 Ref:2	$NoRef$:1Ref:1	-
Wage=10	6	0	$NoRef$:3 Ref:3	-
Wage=15	6	$NoRef$:1 Ref:1	$NoRef$:2 Ref:2	-

Note: $p-values$ for Mann-Withney tests for continuous variables, Fisher-Exacts test for dichotomous variables.

A3 Wage and profit

The first order condition that defines d_W, the optimal transfer for participant i relative to his direct utility function is given by equation A3.1.

$$m'(\pi_i) - \alpha m'(\pi_j) = 0 \tag{A3.1}$$

Define $g = m'(\pi_i) - \alpha m'(\pi_j)$. In equation A3.2, we differentiate g as respect to d and W. We denote differential Δ in order to distinguish it from the transfer d.

$$\begin{aligned}\Delta g = &\frac{\partial \pi_i}{\partial d}\frac{\partial m'(\pi_i)}{\partial \pi_i}\Delta d + \frac{\partial \pi_i}{\partial W}\frac{\partial m'(\pi_i)}{\partial \pi_i}\Delta W \\ &- \alpha\frac{\partial \pi_j}{\partial d}\frac{\partial m'(\pi_j)}{\partial \pi_j}\Delta d - \alpha\frac{\partial \pi_j}{\partial W}\frac{\partial m'(\pi_j)}{\pi_j}\Delta W\end{aligned} \tag{A3.2}$$

We want $\Delta g = 0$ in order for the first order condition to continue holding after changes in d and W. With some omitted steps, $\Delta g = 0$ is found for the relative change of d over W reported in equation A3.3.

$$\frac{\Delta d}{\Delta W} = -\frac{m''(\pi_i) - \alpha m''(\pi_j)}{E(m''(\pi_i) + \alpha m''(\pi_j))} \tag{A3.3}$$

$0 < \frac{\Delta d}{\Delta W} < 1$, which implies an increase in wage for the dictator will be only partly passed on to the recipient. Thus, an increase in wage will increase both i and j's profits.

A4 Continuity and concavity of the utility function

Continuity

The utility function is continuous excepted for transfers $d = \tilde{d}$ where $\tilde{d}$ is such that $\pi_i = \overline{\pi_i}$ and $\pi_j = \overline{\pi_j}$. For this value, it is yet to be proven.

$$\lim_{d \to \tilde{d}^-} u = m(\overline{\pi_i}) + \alpha m(\overline{\pi_j}) + \lambda(m(\pi_i) - m(\overline{\pi_i}))^\beta + \alpha\lambda(m(\pi_j) - m(\overline{\pi_j}))^\beta$$

$$\lim_{d \to \tilde{d}^+} u = m(\overline{\pi_i}) + \alpha m(\overline{\pi_j}) + (m(\pi_i) - m(\overline{\pi_i}))^\gamma + \alpha(m(\pi_j) - m(\overline{\pi_j}))^\gamma$$

It is easy to see that $\lim_{d \to \tilde{d}^-} u = \lim_{d \to \tilde{d}^+} u$: u is continuous at $d = \tilde{d}$.

Concavity

The concavity of the utility function in the gain domain directly follows from the concavity of $m()$ and $\mu()$ in the gain domain. In the loss domain, $\mu()$ is convex, but for values of β close to one, u is still concave. To show it, we study the sign of the second derivative of the utility function with respect to π_i in the loss domain, depicted in Equation A4.1:

$$\begin{aligned} u'' = {} & m''(\pi_i) + \alpha m''(\pi_j) + \lambda\beta m''(\pi_i)(-(m(\pi) - m(\overline{\pi_i}))^{\beta-1} \\ & + \lambda\beta(1-\beta)m'(\pi_i)(-(m(\pi_i) - m(\overline{\pi_i}))^{\beta-2} \\ & + \lambda\alpha\beta m''(\pi_j)(-(\alpha m(\pi_j) - \alpha m(\overline{\pi_j}))^{\beta-1} \\ & + \lambda\alpha\beta(1-\beta)\beta m'(\pi_j)(-(\alpha(m(\pi_j) - \alpha m(\overline{\pi_j}))^{\beta-2} \end{aligned} \tag{A4.1}$$

For values of β close to 1, this expression is negative, and u is concave. Such values of β denote a low diminishing sensitivity in the loss domain and corresponds to the empirical measures. For an overview of the estimates of β in the literature, see Fox and Poldrack (2014).

A5 Details of the calculation for proof 2 & 3

Calculation for proof 2. In order to prove proposition 2, we must show that $m'(\pi_i)\beta(-\lambda)(-1)(m(\overline{\pi_i}) - m(\pi_i))^{\beta-1} - m'(\pi_j)\alpha\beta(\lambda)(m(\overline{\pi_j}) - m(\pi_j))^{\beta-1})$ is positive.

Since $m'(\pi_i) - \alpha m'(\pi_j) = 0$, the former expression is positive when $(m(\overline{\pi_i}) - m(\pi_i))^{\beta-1} > (\alpha m(\overline{\pi_j}) - \alpha m(\pi_j))^{\beta-1}$. Given that $\beta - 1 < 0$, it is equivalent to $(m(\overline{\pi_i}) - m(\pi_i)) < (\alpha m(\overline{\pi_j}) - \alpha m(\pi_j))$. This can be rewritten $m(\pi_i) - \alpha m(\pi_j) > m(\overline{\pi_i}) - \alpha m(\overline{\pi_j}))$.

Remember that $\pi_i > \pi_j$, $\overline{\pi_i} > \overline{\pi_j}$ and $\overline{\pi_i} > \pi_i$. Moreover, by definition, $\pi_i - \pi_j = \overline{\pi_i} - \overline{\pi_j}$. Since $m()$ is concave, the previous implies that the derivative is positive at this point.

Calculation for proof 3. In order to prove proposition 3, we must show that $m'(\pi_i)\gamma(m(\pi_i) - m(\overline{\pi_i}))^{\gamma-1} - m'(\pi_j)\alpha\gamma(\alpha m(\pi_j) - \alpha m(\overline{\pi_j}))^{\gamma-1}$ is negative.

Since $m'(\pi_i) - \alpha m'(\pi_j) = 0$, the former expression is negative when $(m(\pi_i) - m(\overline{\pi_i}))^{\gamma-1} < \alpha(m(\pi_j) - m(\overline{\pi_j}))^{\gamma-1}$. Given that $\gamma - 1 < 0$, it is equivalent to $m(\pi_i) - m(\overline{\pi_i}) > \alpha m(\pi_j) - \alpha m(\overline{\pi_j})$. This can be rewritten $m(\pi_i) - \alpha m(\pi_j) > m(\overline{\pi_i}) - \alpha m(\overline{\pi_j}))$.

Remember that $\pi_i > \pi_j$, $\overline{\pi_i} > \overline{\pi_j}$ and $\overline{\pi_i} < \pi_i$. Moreover, by definition, $\pi_i - \pi_j = \overline{\pi_i} - \overline{\pi_j}$. Since $m()$ is concave, the previous implies that the derivative is negative at this point.

Table A6.1 – Pairwise correlations between responses in the emotion questionnaire.

	Anger	Surprise	Disap.	Joy	Elation	Jealousy	Rage	Frust.	Irrit.	Sad.
Anger	1									
Surprise	0.185**	1								
Disap.	0.594***	-0.018	1							
Joy	-0.207***	0.416***	-0.544***	1						
Elation	-0.055	0.377***	-0.290***	0.650***	1					
Jealousy	0.547***	0.050	0.632***	-0.239***	-0.106	1				
Rage	0.668***	0.071	0.548***	-0.156**	-0.0146	0.642***	1			
Frustration	0.572***	-0.002	0.762***	-0.363***	-0.216***	0.668***	0.564***	1		
Irritation	0.779***	0.054	0.615***	-0.243***	-0.144**	0.672***	0.691***	0.713***	1	
Sadness	0.679***	0.054	0.640***	-0.206***	-0.096	0.677***	0.694***	0.666***	0.772***	1

$N = 190$, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, Bonferroni corrections for multiple comparisons are used.

A6 Analysis of the self-reported emotion data.

Table A6.1 gives the pairwise correlation between responses to the 10 items of the emotion questionnaire. We ran a factor analysis to check whether the questionnaire can be reduced to a small set of factors. This analysis confirms that two factors capture most of the variation in the responses. The first factors is best explained by the negative emotions: Anger, Disappointment, Jealousy, Rage, Frustration, Irritation and Sadness. The second factor is best explained by the positive emotions: Joy, Elation and Surprise.

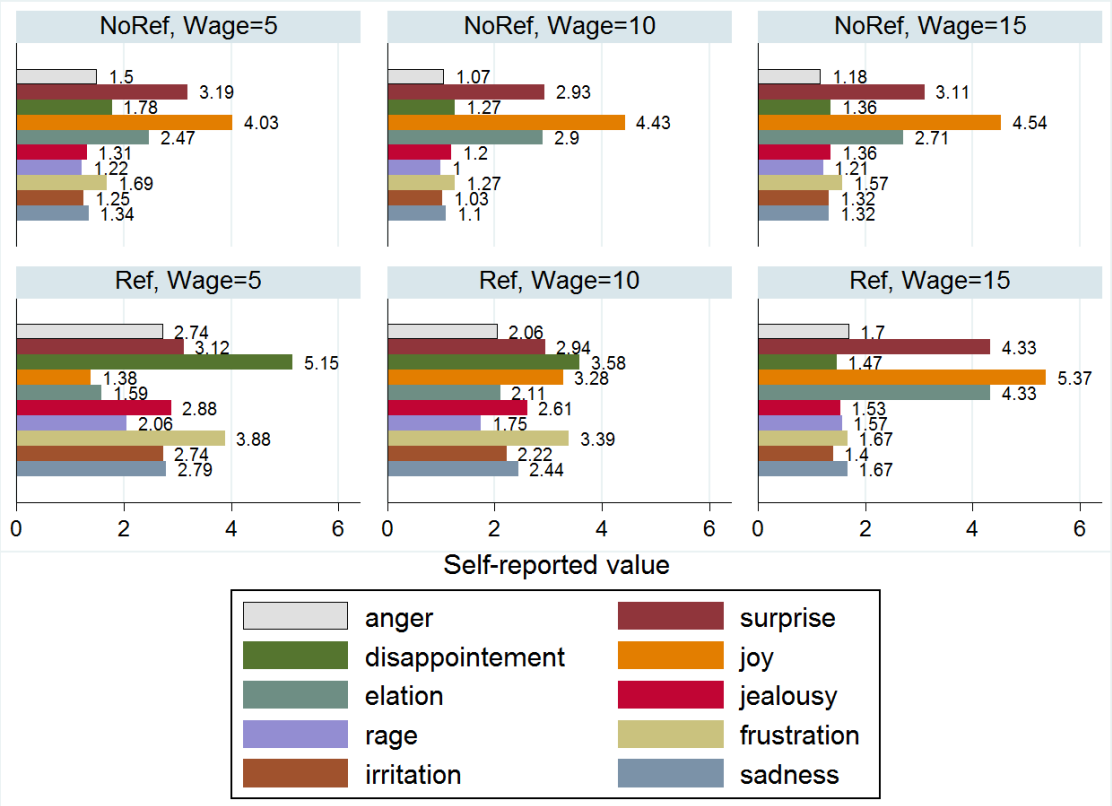


Figure A6.1 – Average responses for each items of the emotion questionnaire, by condition

General Conclusion

The chapters of the present thesis study three social determinants of economic behavior: social preferences, social image and social identity. Understanding the way individuals are influenced by these social determinants has the potential to provide deep insights on a variety of economic questions.

Chapter 1 uses the concept of social identity to study the impact of upward social mobility on trust. The results of this chapter highlight the potential perils of upward mobility, as the promoted tend to trust less both their ingroups and outgroups. On the other hand, we find no evidence that upwardly mobile individuals are discriminated against. In particular, the left-behinds do not seem to be envious of the upwardly mobiles, which is good news.

Taken together, the results indicate that upward mobility comes with some costs. Moreover, the results suggest that upward mobility has some complex effects on the identity of the promoted individuals: they feel “singled-out" by mobility, and define themselves neither by their natural identities, nor by their statuses but rather by a subtle combination of these. An implication is that firms who are willing to favor promotion must accompany it with some measures to avoid that promoted individuals feel isolated. On a more fundamental note, this study shows that a lab induced status that does not exist beyond the time frame of the experiment can undo the identity provided by membership into a salient natural group. An implication might be that assigning statuses

in an homogeneous group whith a common identity can lead to social fragmentation. This indicates how strongly statuses, especially high ones, are used by individuals to define who they are. This echoes previous findings on the effects of status on economic decisions (Ball et al., 2001) and on the costs individuals are willing to bear in order to have a high status (Charness et al., 2014).

Limits and extensions In chapter 1, we tried to take an important problem to the lab in order to have great controls on confounding factors. An obvious limit of this endeavor is that we simplified the phenomenon of upward mobility to a great extent, which might raise questions about the generalizability of our results. For instance, we kept income equal between statuses, whereas in the real life, promotions generally come with greater income. We see our study as a good benchmark to study the effect of upward mobility but some ecologically-valid dimensions of the problem studied could be introduced. Second, we chose natural identities to study upward mobility. Maybe some specific norms or culture partially explain our results and it would be interesting to study upward mobility in groups that vary in terms of norms and culture. We also think that an interesting extension would be to use a similar design to study demotion, which is a great concern of people in western countries nowadays.

Chapter 2 studies the impact of social image on misreporting in a setting in which individual misreports can spillover to the image of the ingroup. We find that social image as a public good feature, as the behavior of one member of a group spills over to the social image of the ingroup. Nevertheless, these spillovers are not internalized and individuals misreport at the expense of their ingroup's image and economic opportunity. This is probably because individuals fail to anticipate that their behavior impacts their ingroups' economic opportunities. Our results show that the individual concern for social image does not translate to group settings, because of mispecified beliefs.

The results of chapter 2 provide insights on the way group members contribute to the image of the group, an important question that has been barely studied in economics (Tirole, 1996; Huck and Lünser, 2010). The results from chapter 2 emphasize that groups face public good problems when it comes to building a common social image. As for any public good, there might be under provision especially if individuals do not internalize the impact of their behavior on the image of the group, as is the case in our experiment. This in turn might impede the economic prosperity of social groups that cannot implement solutions to improve the behavior of the members.

Limits and extensions In chapter 2, we test the effect of image spillovers with a simple and neat design. It is possible that the starkness of the design explains why individuals do not internalize the effect of their behavior on the image of the group, as they have no learning opportunities. An interesting extension of our work could be to study specifically under which conditions and how effectively individuals learn that their behavior is detrimental to the social image of their group. Also, we implemented our experiments with friends, a very specific kind of social group. It would be interesting to see if our results hold with different types of groups. For instance, we could replicate our study with members of stereotyped social groups because they might have a different understanding of the public good features of the image of the group, and they could have specific responses to it.

Chapter 3 studies the impact of counter-factual payoffs on generosity. We find that participants who get a lower payoff that the one they could expect tend to transfer less, and those who get a higher payoff that the one they could expect tend to transfer more. We find that emotions do not drive these results while our simple model of reference-dependent generosity has a good predictive power. The results suggest that participants feel poorer (richer) when they get less (more) money than expected and their transfers are responsive to it.

The results of chapter 3 indicates that generosity is not driven only by one's payoff (or relative payoff) but also by how one evaluates this payoff. More generally, these results contribute to the literature showing that social preferences are not absolute but rather depend on the context of the decisions. Subtle features of the context can for instance provide individuals with an excuse not to give (Dana et al., 2006), can change the perceptions about what is the norm (List, 2007; Bardsley, 2008) or convince people that they are entitled not to give (Cherry et al., 2002). Our contextual manipulation, varying the information about how much participants could have earned, impacted the strength of the concern for the recipient's payoff. The sensitivity of social preferences to the context has raised doubts on the internal validity of the economic concept and its predictive power (Wilson, 2010; Binmore and Shaked, 2010) as well as on the external validity of its measurement (Levitt and List, 2007). In contrast, we believe that the meticulous study of the contextual details that impact social preferences is a fertile endeavour (Falk and Heckman, 2009). For instance, List (2007) and Bardsley (2008) find that adding a "take" option in the dictator game reduces transfers. This should not lead us to consider that altruism is an artifact of the experimental procedures, but should rather promote the idea that context-specific norms matter. This should in turn spark the interest in what determines norms and how norms affect social preferences.

Limits and extensions In chapter 3 we develop a model that provides good predictions, but its internal validity hinges on the assumptions we made about the reference point. We made the reasonable assumption that the reference point lies in-between the two extreme values of the wage. It is possible that participants formulate their reference differently. More importantly, it is possible that our aggregate results hide a large individual heterogeneity. For instance, among the participants who received the medium wage, some may consider it as a loss because their reference is the high wage and some others may consider it as a gain because their reference is the low wage. Our data provides no insights on this question. A possible extension of our research could

be to vary the number of potential levels of wage within participants to study more precisely how participants evaluate their wage and how it impacts subsequent transfers.

In this dissertation, we study social identity, social image concerns and social preferences independently in three separate chapters. This approach is analytic, and is not due to the belief that these motivations are independent. Indeed, they shape individual economic behavior together and interact. For instance, chapter 1 uses a trust game to reveal the effect of upward mobility on trust. Behavior in trust games are determined in part by social preferences, and as such the results from chapter 1 contribute to the literature studying the impact of social identity on social preferences (Bowles and Gintis, 2004; Goette et al., 2006; Charness et al., 2007; Chen and Li, 2009). Similarly, the analysis proposed in chapter 2 focuses on how the social image of groups is built, and it is likely that social preferences and social identity also matter in the decision to contribute to the image of the group. For instance, the concern for an ingroup's payoff is integrated in our analysis as a motivation for individuals to internalize the impact of their behavior on the image of the group. A greater concern should push individuals to act better in order to preserve the social image and the economic opportunities of the ingroup. Similarly, social identity may also matter. Individuals identify with a group if it can help them perceive themselves positively. Individuals should not identify with social groups with poor social image, and as a consequence should not care much about the social image of the group. In the last chapter, the study focuses on social preferences and more specifically generosity. The focus is to test how transfers in a dictator game reacts to gains and losses relative to a reference, and such transfers have been shown to help individuals improve their self-concept (Bénabou and Tirole, 2006) and their social image (Andreoni and Bernheim, 2009; Andreoni et al., 2017). It is possible that our treatment effects are driven by the fact that losses and gains change the concern for self-esteem and for social-esteem.

Bibliography

Andreoni, J. and B. D. Bernheim (2009). Social image and the 50-50 norm: A theoretical and experimental analysis of audience effects. *Econometrica 77*(5), 1607–1636.

Andreoni, J., J. M. Rao, and H. Trachtman (2017). Avoiding the ask: A field experiment on altruism, empathy, and charitable giving. *Journal of Political Economy 125*(3), 625–653.

Ball, S., C. C. Eckel, P. Grossman, and W. Zame (2001). Status in markets. *The Quarterly Journal of Economics 116*(1), 161–188.

Bardsley, N. (2008, Jun). Dictator game giving: altruism or artefact? *Experimental Economics 11*(2), 122–133.

Bénabou, R. and J. Tirole (2006, December). Incentives and prosocial behavior. *American Economic Review 96*(5), 1652–1678.

Binmore, K. and A. Shaked (2010). Experimental economics: Where next? *Journal of Economic Behavior Organization 73*(1), 87 – 100. On the Methodology of Experimental Economics.

Bowles, S. and H. Gintis (2004). Persistent parochialism: trust and exclusion in ethnic networks. *Journal of Economic Behavior & Organization 55*(1), 1–23.

Charness, G., D. Masclet, and M. C. Villeval (2014). The dark side of competition for status. *Management Science 60*(1), 38–55.

Charness, G., L. Rigotti, and A. Rustichini (2007). Individual behavior and group membership. *American Economic Review 97*(4), 1340–1352.

Chen, Y. and S. X. Li (2009). Group identity and social preferences. *The American Economic Review 99*(1), 431–457.

Cherry, T. L., P. Frykblom, and J. F. Shogren (2002). Hardnose the dictator. *The American Economic Review 92*(4), 1218–1221.

Dana, J., R. a. Weber, and J. X. Kuang (2006, sep). Exploiting moral wiggle room: experiments demonstrating an illusory preference for fairness. *Economic Theory 33*(1), 67–80.

Falk, A. and J. J. Heckman (2009). Lab experiments are a major source of knowledge in the social sciences. *science 326*(5952), 535–538.

Goette, L., D. Huffman, and S. Meier (2006). The impact of group membership on cooperation and norm enforcement: Evidence using random assignment to real social groups. *American Economic Review 96*(2), 212–216.

Huck, S. and G. K. Lünser (2010). Group reputations: An experimental foray. *Journal of Economic Behavior & Organization 73*(2), 153–157.

Levitt, S. D. and J. A. List (2007). What do laboratory experiments measuring social preferences reveal about the real world? *Journal of Economic perspectives 21*(2), 153–174.

List, J. A. (2007). On the Interpretation of Giving in Dictator Games. *Journal of Political Economy 115*(3), 482–493.

Tirole, J. (1996). A theory of collective reputations (with applications to the persistence of corruption and to firm quality). *The Review of Economic Studies 63*(1), 1–22.

Wilson, B. J. (2010). Social preferences arenâ€™t preferences. *Journal of Economic Behavior & Organization 73*(1), 77–82.

Essays on the economics of social identity, social preferences and social image.

Rémi SUCHON

Abstract

The present dissertation studies three social determinants of economic decisions: Social Identity, Social Image, and Social preferences. The first chapter reports on an experiment testing the effect of upward social mobility on interpersonal trust. Individuals are characterized both by a natural group identity and by a status awarded by means of relative performance in a task in which natural identities strongly predict performance. Upward mobility is characterized by the access to the high status of individuals belonging to the natural group associated with a lower expected performance. We find that socially mobile individuals trust less than those who are not socially mobile, both when the trustee belongs to the same natural group or to the other natural group. In contrast, upward mobility does not affect trustworthiness. We find no evidence that interacting with an upwardly mobile individual impacts trust or trustworthiness. In the second chapter, we test whether individuals internalize the effects of their behavior on the social image of their group. In our experiment, we recruit pairs of real-life friends and study whether misreporting decreases when it may have negative spillovers on the image of the friend. We find that participants hurt their friends' social image by misreporting: external observers update their beliefs and rightfully expect that a participant whose friend misreported is likely to misreport himself. However, participants misreport as often when their behavior can hurt the friend's image as when it cannot, even though hurting their friends' image reduces their own monetary gains. Our interpretation is that they underestimate the impact of their behavior on external observers' beliefs about their friends. Our results show that, even in our case where group membership is salient, groups might have difficulties building a good image. The good news is that external observers may use image spillovers to update their beliefs and interact with members of groups more efficiently. In the third chapter, we experimentally test whether the salience of counter-factual payoffs impacts generosity. Participants first perform a real-effort task for a fixed wage, and then play a dictator game. Between conditions, we vary the level and the timing of the revelation of the wage. In some conditions, participants know the wage before the real effort task, and are not informed of the other potential levels. In some other conditions, they are informed of the distribution of the wages before the real effort task, but the actual wage is only revealed afterward. Our hypothesis is that participants in the latter conditions evaluate their actual wage relative to the other potential levels, which in turns impact their transfers in the subsequent dictator game. The results support this hypothesis: participants who get the high wage tend to transfer more when they are informed of the other potential levels than when they are not. Symmetrically, participants who get the low wage tend to transfer less when they are informed of the other potential levels than when they are not.

Keywords: Social Preferences; Social Identity; Social Image; Experiment

Essais en économie de l'identité sociale, des préférences sociales et de l'image sociale.

Rémi SUCHON

Résumé

Cette thèse porte sur trois déterminants sociaux des décisions économiques : l'identité sociale, l'image sociale et les préférences sociales. Le premier chapitre rend compte d'une expérience visant à tester l'effet de la mobilité sociale ascendante sur la confiance interpersonnelle. Les individus sont caractérisés à la fois par une identité de groupe naturelle et par un statut attribué au moyen de leur performance relative dans une tâche dans laquelle les identités naturelles prédisent fortement la performance. La mobilité ascendante se caractérise par l'accès au statut élevé des individus appartenant au groupe naturel associé à une performance attendue inférieure. Nous constatons que les personnes socialement mobiles font moins confiance que celles qui ne sont pas socialement mobiles, à la fois lorsque l'autre individu appartient au même groupe naturel ou à un autre groupe naturel. En revanche, la mobilité ascendante n'affecte pas la fiabilité. Rien n'indique que l'interaction avec une personne mobile a une incidence sur la confiance ou la fiabilité. Dans le deuxième chapitre, nous testons si les individus intériorisent les effets de leur comportement sur l'image sociale de leur groupe. Dans notre expérience, nous recrutons des paires d'amis et étudions si la frÃ©quence des fausses déclarations diminue quand cela peut avoir des retombées négatives sur l'image de l'ami. Nous constatons que les participants nuisent à l'image sociale de leurs amis en faisant de fausses déclarations : les observateurs externes mettent à jour leurs croyances et s'attendent à juste titre à ce qu'un participant dont l'ami a fait de fausses déclarations soit susceptible de faire de même. Cependant, les participants font autant de fausses déclarations quand leur comportement peut nuire à l'image de leur ami que quand il ne le peut pas, même si le fait de nuire à l'image de leurs amis réduit leurs propres gains monétaires. Notre interprétation est qu'ils sous-estiment l'impact de leur comportement sur les croyances des observateurs externes au sujet de leurs amis. Nos résultats montrent que, même dans notre cas où l'appartenance à un groupe est évidente, les groupes peuvent avoir des difficultés à établir une bonne image. Les observateurs externes peuvent utiliser les retombées d'images pour mettre à jour leurs croyances et interagir plus efficacement avec les membres des groupes. Dans le troisième chapitre, nous examinons expérimentalement si la sailliance des gains contre-factuels a une incidence sur la générosité. Les participants exécutent d'abord une tâche d'effort réel pour un salaire fixe, puis jouent au jeu du dictateur. Entre les conditions, nous faisons varier le niveau et le moment de la révélation du salaire. Dans certaines conditions, les participants connaissent le salaire avant la tâche d'effort réel et ne sont pas informés des autres niveaux potentiels. Dans d'autres conditions, ils sont informés de la répartition des salaires avant la tâche d'effort réel, mais le salaire effectif n'est révélé qu'ensuite. Notre hypothèse est que les participants, dans ces dernières conditions, évaluent leur salaire effectif en comparaison des autres niveaux potentiels, ce qui, à son tour, influe sur les transferts dans le jeu dictateur. Les résultats corroborent cette hypothèse : les participants qui obtiennent un salaire élevé ont tendance à transférer davantage lorsqu'ils sont informés des autres niveaux potentiels que lorsqu'ils ne le sont pas. Symétriquement, les participants qui reÃ§oivent le plus bas salaire ont tendance à transférer moins quand ils sont informés des autres niveaux potentiels que quand ils ne le sont pas.

Mots Clés: Préferences sociales; Identité sociale; Image sociale; Expérience