Sede Amministrativa Università degli Studi di Padova
Dipartimento di Scienze Economiche e Aziendali "Marco Fanno"

SCUOLA DI DOTTORATO DI RICERCA IN
ECONOMIA E MANAGEMENT
CICLO XXV

Lies, Incentives and Self-confidence

Direttore della Scuola: Ch.mo Prof. Giorgio Brunello
Supervisore: Ch.mo Prof. Antonio Nicolò

Dottoranda: Valeria Maggian

Risate, signor Meis,
di tutte le vane, stupide afflizioni che esso ci ha procurate,
di tutte le ombre, di tutti i fantasmi ambiziosi e strani
che ci fece sorgere innanzi e intorno,
della paura che c'ispirò!

- Il fu Mattia Pascal -

Luigi Pirandello

To my mother and my father

To Daniela and Attilio

Acknowledgements

I’ve dreamt this moment for a long time (among other nightmares related to the Phd). Now I can finally thank all the people who have guided, sustained and helped me during these years.

I am deeply grateful to my advisor, Antonio Nicolò. In addition to guiding and supporting me during this Phd path, he has spread his real passion for research and teaching through his empathy and sense of humor.

Special thanks go to Marie Claire Villeval. Her advising in Lyon has been enormously beneficial to me, both from professional and personal point of view. As one chapter of my thesis is related to women’s underconfidence, she represents to me a clear example of how an exception to such aggregate evidence can make the difference, in real life.

I am indebted to my coauthor Natalia Montinari, I am benefiting a lot from all our stimulating discussions, from the advices received and from your many natural human qualities.

I also want to thank Marco Piovesan and Alessandro Bucciol, who firstly instilled in me the pleasure of doing research and who have supported and encouraged me, especially in the first steps of my career.

For their kind assistance with writing letters, giving wise advice, helping with various applications, and so on, I wish to thank in addition Luca Corazzini, Sthépane Robin, Fabio Manenti, Paola Valbonesi and Efrem Castelnuovo.

Thank you to my great colleagues Gine, Ambra, Elena, Gabriel, Valentina, Giulia, Anna, Marco, Roberto and Andrea for making the Economics Department in Padova a great place to work. I am especially grateful to Michele Fabrizi, Paolo Roberti, Michele Costola, Pablo Torija Jemenez, Diana Gaspari and Petro Bonetti: I’m still laughing when thinking about all the unforgettable moments we had.
Thanks to Cédric for helping me during the first days in Lyon, and I’m also grate for having met so many friendly people in these years. Thank you, Michela, for the pleasant stay in Firenze. Thanks to Dasha, Nicolas, Wenceslao, Maria, Tushi, Emanuele and Natalia for the funny days in Trento and Julie, Ranoua, Julie, Thierry and Katerine for your smiles in Lyon.

It would have been tremendously harder and dourer to face this experience without the priceless presence of my friends Giulia G., Giulia R., Giulia D.V. and Daniela. Thank you also to my oldest friends, Valentina, Cristina and Alice who are always there for me.

Thank you, Attilio, for all the moments we have shared.

Most importantly, I want to thank Maria, my mother, and Pietro, my father, for much more than supporting me during the Phd adventure. Thank you for letting me to be myself, for making me experience the true sense of respect and love.

Valeria Maggian

Contents

List of Figures

Chapter 1

Chapter 2

List of Tables

Chapter 2

Introduction

Lies, Incentives and Self-Confidence

The present thesis is composed by three chapters, each of them making contributions to three distinct topics in behavioral Economics. The chapters can thus be read independently from each other. The first chapter concerns an experimental analysis which aim is to examine the development of social preferences with respect to age and how they are related with lying behavior of children. The second chapter investigates the role of reciprocity in exacerbating inefficient and opportunistic behavior in hierarchical organizations, when conflict of interests between its members arises at different levels. Finally, the third chapter contains a theoretical model explaining the emergence of the gender gap in top job positions in line with recent evidence provided by experimental Economics. While both the first and the second chapter in this thesis present experimental evidence, the experimental methods differ across them. In particular, in the first chapter we report evidence from an artefactual field experiment while the second chapter presents a conventional framed laboratory experiment[1]. By using experiments, we are able to control and to impose exogenous variation to the environment where people make decisions. The third chapter applies a theoretical approach, with the model based on recent experimental evidence on the research topic of interest. Experimental Economics represents a useful contributor to Economic theory, strengthening its predictions or providing new insights for future development (Falk and Heckman, 2009). In the first two chapters of the present thesis, experiments allow us to observe and investigate the development of lying behavior during childhood and the emergence of an opportunistic conduct in the job environment, observations which are difficult to detect with empirical data. Finally, in the third chapter we translate the recent experimental research on gender differences in a more formal and tractable manner, through a stylized model.

The first chapter of the thesis experimentally examines how the interaction between lying aversion and social preferences affects children's behavior when playing a modified version of the dictator game. In the last decade, many experimental studies in Economics have investigated the role of other regarding preferences in determining the behavior of individuals:

[1] According to the taxonomy of Harrison and List (2004).

these empirical findings have confirmed that individuals are not just interested in maximizing their own welfare but, with some limitations, they have a concern for the welfare of the others (Henrich *et al.,* 2004, Fehr and Schmidt, 2006). Paralleling to this stream of investigation, the lying behavior of individuals has been objective of interest of many experimental studies (Gneezy, 2005; Mazar *et al.*, 2008). However, it is still a matter of discussion whether these features are innate or if they evolve over time, and how different incentives and contexts affects their development. In order to answer these questions it is interesting to analyze the behavior of children when facing relevant economic decisions in a controlled environment. In particular, in a sample of 637 children, aged between 7 and 14, we find strong aversion to lying at all ages, and especially among females. We find that while children become more concerned about other's welfare as aging, they are not prone to violate moral norms in order to implement their other-regarding preferences. On the contrary, lying is more likely to be determined by selfish and envy motives, at all ages.

The second chapter investigates the emergence of a dark side of reciprocity in the working environment. Reciprocity has been shown to be a desirable feature in organizations, reducing the costs of aligning interests between the principal and the agent. Both laboratory and field experiments (see Fehr and Falk (2008) for an overview of recent results), have provided evidence that workers respond to generous wage levels by exerting above minimal effort, confirming the gift-exchange hypothesis first formulated by Akerlof (1982). In hierarchical organizations, however, conflict of interests between its members may emerge at multiple levels: in such a situation, reciprocity exacerbate, rather than alleviate, the negative effects of members' misalignment of objectives. The second chapter of this thesis consists in a laboratory experiment which aim is to analyze the selection and effort distortions resulting from agents persecuting their personal interest by taking advantage of workers' reciprocal concerns. The results show that sharing part of the three-level hierarchical organization's profit with workers, rather than only with agents, impedes the latter to exploit their powerful position at the disadvantage of the organization.

The third chapter is about the role of self-confidence in determining females' underrepresentation in high skilled occupations. Relying on recent experimental evidence which explains the occupational gender gap as the result of different preferences and attitudes of men and women with respect to risk, ambition, self-confidence and willingness to compete

(Gneezy *et al.*, 2003; Niederle and Vesterlund, 2007; Datta Gupta *et al.*, 2013), we derive a model to explain the emergence of gender segregation in the labor market as a result of females' biased beliefs regarding their ranking position with respect to other (male) candidates, when abilities are equally distributed among them. In particular, in our model, women self-select into low-skilled occupations according to their (mis)perceptions about their opportunity to be successfully recruited when competing for better positions. The third chapter further illustrates the powerful impact of affirmative actions in restoring efficiency in the job matching equilibrium between firms and workers. Indeed, in line with recent experimental evidence (Balafoutas and Sutter, 2012; Niederle *et al.*, 2013), we provide a theoretical foundation to sustain the importance of implementing calibrated gender quota in order to restore the efficiency of job matching between high skilled firms and candidates, by encouraging skilled women to enter competition for top job market positions and thus increasing the diversity of qualified applicants.

References

Akerlof, G. Labor Contracts as Partial Gift-Exchange. *Quarterly Journal of Economics* **97**, 543-569 (1982).

Balafoutas, L. and Sutther, M. Affirmative action policies promote women and do not harm efficiency in the laboratory. *Science* **335**, 579-582 (2012).

Datta Gupta, N., Poulsen, A., and Villeval, M. C. Gender matching and competitiveness: experimental evidence. *Economic Inquiry* **51**, 816-835 (2013).

Falk, A. and Heckman, J. Lab Experiments Are a Major Source of Knowledge in the Social Sciences. *Science* **326**(5952), 535-538 (2009).

Fehr, E. and Falk, A. *Reciprocity in Experimental Markets* (Handbook of Experimental Economic Results, C.R. Plott and V.L. Smith, eds, 2008).

Fehr, E., and Schmidt, K.M. *The economics of fairness, reciprocity, and altruism—Experimental evidence and new theories* (In: Kolm, S., Ythier, J.M. (Eds.), Handbook on the Economics of Giving, Reciprocity and Altruism, vol. 1. North-Holland, Amsterdam, 2006).

Gneezy, U. Deception: The Role of Consequences. *American Economic Review* **95**, 384-394 (2005).

Gneezy, U., Niederle, M., and Rustichini, A. (2003). Performance in competitive environments: Gender differences. *The Quarterly Journal of Economics* **118**, 1049-1074 (2003).

Harrison, G.W., and List, J.A. Field Experiments. *The Journal of Economic Literature* **42**(4), 1009-1055 (2004).

Henrich, J., Boyd, R., Bowles, S., Gintis, H., Fehr, E. and Camerer, C. *Foundations of human sociality: Economic experiments and ethnographic evidence from fifteen small-scale societies* (Oxford University Press, 2004).

Mazar, N., Amir, O., and Ariely, D. The dishonesty of honest people: a theory of self-concept maintenance. *Journal of marketing research* **45**, 633-644 (2008).

Niederle, M. and Vesterlund, L. Do women shy away from competition? do men compete too much? *Quarterly Journal of Economics* **122,** 1067-1101 (2007).

Niederle, M., Segal, C., and Vesterlund, L. How costly is diversity? Affirmative action in light of gender differences in competitiveness. *Management Science* **59**, 1-16 (2013).

Introduzione (in Italian)

Lies, Incentives and Self-Confidence

La presente tesi è composta da tre capitoli, ciascuno dei quali facenti riferimento a tre distinti ambiti di ricerca dell'Economia comportamentale. Ciascun capitolo può quindi essere letto in maniera indipendente rispetto agli altri. Il primo lavoro consiste in un'analisi sperimentale volta ad analizzare lo sviluppo di comportamenti disonesti condizionatamente all'emergere di preferenze sociali nei bambini, in diverse fasce d'età e per genere. Il secondo lavoro analizza il ruolo della reciprocità nell'esacerbare, invece che nell'alleviare, l'emergere di comportamenti opportunistici e inefficienti all'interno di organizzazioni di tipo gerarchico, nelle quali i conflitti di interessi si evidenziano su più livelli. Infine, il terzo capitolo consiste in un modello teorico che intende spiegare il presente divario tra uomini e donne nelle posizioni di management basando le proprie assunzioni sulle ultime evidenze sperimentali. Nonostante i primi due capitoli siano entrambi basati su dati sperimentali, il metodo applicato è differente. In particolare, mentre il primo capitolo consiste in un field experiment realizzato con i bambini all'interno delle scuole, il secondo capitolo si basa su un tradizionale esperimento in laboratorio. Attraverso l'economia sperimentale è possibile controllare l'ambiente in cui gli individui prendono le loro decisioni, imponendo delle variazioni esogene sulla variabile di interesse. Il terzo capitolo utilizza un approccio teorico, basando le proprie assunzioni sui recenti sviluppi offerti dall'Economia comportamentale e sperimentale. L'Economia sperimentale rappresenta un importante strumento della teoria economica, ne rafforza i risultati e permette di creare nuove basi per migliorarne il potere predittivo (Falk and Heckman, 2009). Nei primi due capitoli della tesi la realizzazione degli esperimenti ci permette di analizzare lo sviluppo di comportamenti disonesti rispetto all'età e l'emergere di comportamenti opportunistici nell'ambiente di lavoro, entrambi comportamenti difficili da osservare con l'utilizzo di altri metodi di ricerca differenti dall'esperimento. Infine, nell'ultimo capitolo, le recenti evidenze offerte dall'Economia sperimentale relativamente alle differenti attitudini di uomini e donne sono formalizzate in un modello teorico al fine di spiegare il divario occupazionale tra i due sessi.

Il primo capitolo della tesi ha come obiettivo l'analisi dello sviluppo della disonestà nei bambini, condizionatamente allo sviluppo delle loro preferenze sociali. Tale analisi é realizzata tramite un esperimento basato su una versione modificata del Dictator Game. Negli ultimi decenni, molti lavori sperimentali si sono concentrati sull'analisi del ruolo delle preferenze sociali nell'influenzare le decisioni degli individui: i risultati dimostrano che gli agenti economici non sono interessati esclusivamente alla massimizzazione del loro interesse personale ma, seppure con alcune limitazioni, tengono in considerazione anche gli effetti che le loro decisioni avranno sul benessere di altri indivdui (Henrich *et al.*, 2004, Fehr and Schmidt, 2006). Parallelamente a questa area di ricerca, l'economia sperimentale si è concentrata sulle condizioni e i contesti che determinano comportamenti disonesti negli individui (Gneezy, 2005; Mazar *et al.*, 2008). Eppure, ancora non è chiaro se tali evidenze siano il risultato del processo di socializzazione degli individui o siano caratteristiche innate. Per tale motivo è interessante esaminare in un ambiente controllato come diversi incentivi e contesti influenzino l'evoluzione di tali comportamenti nei bambini, in situazioni economicamente rilevanti. In particolare, nell'esperimento realizzato con un campione di 637 bambini, tra i 7 e i 14 anni, abbiamo rilevato una forte avversione a mentire a tutte le età, specialmente tra le bambine. Nonostante al crescere dell'età i bambini siano più propensi a valutare l'effetto delle loro scelte sul payoff altrui, essi non sono più propensi a mentire per apportare benefici monetari ad altri loro pari. Al contrario, la disonestà è maggiore tra i bambini che esprimono preferenze di tipo egoista, indipendentemente dall'età.

Il secondo capitolo analizza l'emergere di un "lato oscuro" della reciprocità nell'ambiente di lavoro. La reciprocità è solitamente considerata come una caratteristica desiderabile all'interno delle organizzazioni, in quanto riduce il costo di riallineare gli interessi divergenti tra il principale e l'agente. Esperimenti recenti, sia realizzati in laboratorio che nel mondo reale, hanno dimostrato che i lavoratori reciprocano il loro datore di lavoro lavorando più di quanto ci si aspetterebbe quando viene loro offerto uno stipendio al di sopra del minimo accettabile, confermando l'ipotesi del *gift-exchange* formulata inizialmente da Akerlof (1982). Tuttavia, nel secondo capitolo dimostriamo come nelle organizzazioni gerarchiche gli interessi dei membri che le compongono possano divergere su più livelli: in tale situazione la reciprocità esaspera, invece di alleviare, le inefficienze risultanti da tali conflitti di intesse. Il secondo capitolo di questa tesi analizza, attraverso un esperimento in laboratorio, le distorsioni sia sulla selezione dei lavoratori che sulla loro produttività come conseguenza del comportamento opportunista degli agenti i quali, sfruttando la reciprocità dei lavoratori, perseguono i propri interessi personali a danno dell'organizzazione in cui sono assunti. I risultati dell'esperimento dimostrano inoltre che condividere una piccola parte dei profitti dell'organizzazione anche con i lavoratori impedisce agli agenti di sfruttare la loro posizione di potere in modo illegittimo.

Il terzo capitolo è relativo al ruolo della fiducia in se stessi nel determinare il limitato numero di donne in posizioni lavorative di alto livello. Recenti risultati sperimentali mostrano come la sotto-rappresentanza delle donne nel management delle organizzazioni possa dipendere dalle loro diverse attitudini e preferenze rispetto al rischio, all'ambizione, alla fiducia in se stessi e alla volontà di competizione rispetto agli uomini (Gneezy *et al.,* 2003; Niederle and Vesterlund, 2007; Datta Gupta *et al.*, 2013). Il terzo capitolo della tesi presenta quindi un modello teorico che spiega l'emergere della segregazione occupazionale come il risultato delle percezioni erronee delle donne relativamente alla loro (inferiore) abilità rispetto agli uomini, quando invece le abilità sono equamente distribuite tra i due sessi. Il particolare, il modello dimostra come le donne si auto-selezionino in lavori mediocri come conseguenza alla loro (errata) convinzione di non essere all'altezza della selezione quando sono in competizione con gli uomini per posizioni di più alto livello. Al fine di restaurare l'efficienza dell'incontro tra domanda e offerta nel mondo del lavoro, il modello dimostra quindi l'importanza di implementare delle quote che garantiscano l'accesso delle donne nelle suddette posizioni lavorative. Infatti, in linea con quanto recentemente dimostrato in ambito sperimentale (Balafoutas and Sutter, 2012; Niederle *et al*., 2013), il terzo capitolo spiega come tali politiche possano positivamente influenzare il mercato del lavoro, aumentando la diversità dei candidati ma senza intaccarne l'efficienza.

Chapter 1

Social preferences and lying aversion in children

with Marie Claire Villeval

1.1 Introduction

The anthropological reductionism applied by economists for several decades in the ideal figure of *homo oeconomicus*, has been challenged from at least two perspectives. First, extensive research has established that individuals make conscious decisions revealing other-regarding preferences towards strangers[1-5], making cooperation and collective actions achievable[6-7]. Second, contrary to the notion that individuals rationally violate moral norms provided this brings marginal net benefits, recent research provides evidence that most people value honesty[8-10]. Honesty enhances mutual trust, conditions the efficacy of policies[11-12], and is fundamental for preserving human dignity, according to the Kantian categorical imperative[13]. However, pursuing a moral conduct may sometimes conflict with the desire to preserve or improve others' welfare. A consequentialist approach suggests that whether an action is considered morally acceptable depends on its outcome. Thus, the decision to lie is sensitive to the available incentives, to the consequences of the lie on others' well-being[14-17], and to the environment that conditions what is considered as morally acceptable in a given situation[8,10,18]. When pursuit of other-regarding preferences implies lying, individuals face a moral dilemma and how they solve it remains an open question[16].

Analyzing such a dilemma is extremely relevant in childhood, a crucial phase in the process of development of both other-regarding preferences[19-25] and moral reasoning[26]. From the early years parents instill their children the value of honesty while, by entering the schools, they also develop their understanding of others as intentional beings motivated by desires that may come into conflict with their own. As aging, most of the children learn to balance other-regarding acts with self-interest[20-26], but what happens when the pursuit of social preferences implies lying remains unexplored[27].

The development of moral reasoning has been a focus of investigation in psychology[28-29] but not in economics[30]. Children's understanding of the opportunity to not tell the truth in order to respect etiquette, to pursue personal interest, or to avoid punishment emerges as early as three years of age[31-35]. Deceitful behavior evolves during school years as children develop executive control functions and theory of mind[36-37], and the ability to infer correctly which social and moral norms may be violated if they lie[38-39]. However, most studies investigate children's understanding of social rules[40] but disregard how the decision to lie is affected by its consequences on others' welfare. A priori, there are two main plausible developmental pathways when considering how children cope with such a dilemma, as aging. On one hand, since adults are sensitive to the consequences of their lies on other's payoffs[15-17] we might expect that older children, being more likely to care about other's welfare as they age[19-25], are also more likely to lie to benefit others. On the other hand, adults are sometimes unconditionally lie-averse so that we might expect the internalization of the value of honesty to progressively strengthen in childhood. In this paper we analyze whether children's deceitful behavior depends on their other-regarding preferences when economic incentives are present. We conducted an experiment with 637 Italian children (326 females and 311 males) in three age groups, from middle childhood (7-8 and 9-10 years old) to early adolescence (11 and 14 years old). We gave children the opportunity to lie in order to achieve their preferred outcome, making them conscious that their choice would influence both their own payoff and their partner's welfare.

We study whether having a preference for a particular allocation of resources (i.e. an other-regarding or a self-regarding allocation) affects children's willingness to lie to achieve it,according to their age and gender[41-42].

1.2Experimental design

To assess children's other-regarding preferences and willingness to lie, we designed a two-part experiment using the z-Tree software[43]. The experiment is engaged in privately by each child, in a separate room (see Appendix). In the first part, the child has to decide how to allocate a certain amount of points (to be exchanged for pencils, stickers, etc.) between himself and a randomly assigned anonymous classmate. Two options are available: one option gives both children the same number of points; the alternative, depending on the treatment, assigns the child more or less points than his/her partner. The child's choice indicates his/her social orientation. In the second part, each child randomly draws one of the same two options and is asked to report this privately observed random allocation. However, there is the opportunity to misreport which option is observed in order that the child gets his/her preferred one.

Each child participated in one of three treatments. One allocation option is common to all three treatments, the alternative in all three cases differs across them. The common option enables an egalitarian distribution of 5 points to both the decision-maker and his/her partner. The equal share is chosen as the benchmark in the experiment since egalitarianism is a major driver of human actions[4-6]. The alternative option creates advantageous or disadvantageous inequality. In the Selfishness treatment, by choosing the (7,3) alternative option the child increases his/her own payoff by decreasing that of his/her partner. In the Efficiency treatment, he/she can increase the partner's payoff at no cost to him/herself by choosing the (5,7) alternative option. Finally, in the Altruism treatment, the child can increase his/her partner's payoff by choosing the (3,7) option, but this reduces his/her own payoff. Advantageous inequality aversion is observed in the choice of (5,5) instead of (7,3) in the Selfishness treatment. Evidence of altruism is found if the (3,7) option is preferred in the Altruism treatment. Finally, efficiency concerns are identified when the percentage of children choosing the (5,7) option in the Efficiency treatment is significantly above 50%. In this treatment, evidence that significantly more than 50% of children choose the (5,5) option may be due to envy or inequality aversion. If inequality aversion drives children's behavior then we would expect a similar percentage to choose the (5-5) option in the Selfishness and in the Efficiency Treatments.

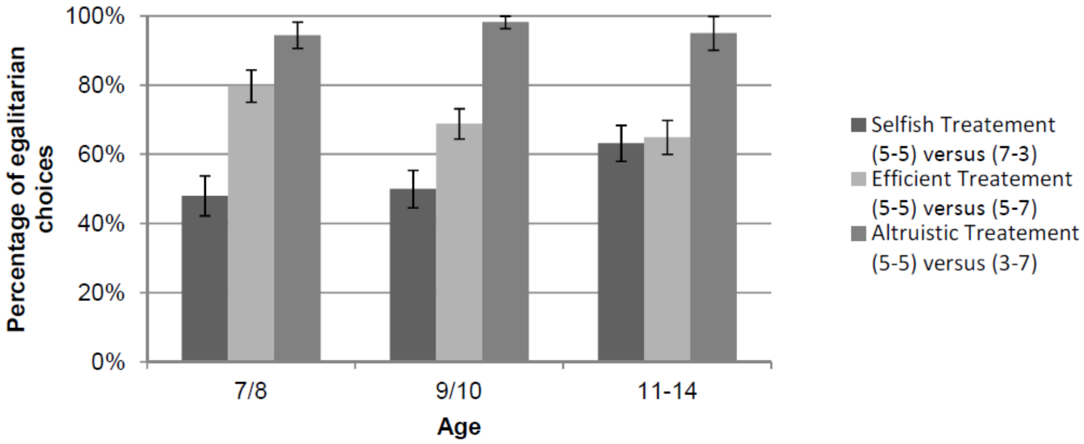
Once the child has chosen his/her preferred allocation in the first part, he/she has to press a button on the computer screen to make a shape - a sun or a star - appear (with the same probability). Each shape corresponds to one of the two options used in the first part. The child is asked to tick on a reporting sheet which shape was privately observed (see Appendix, Fig. A4). Depending on which shape is reported, the egalitarian or the alternative option is actually implemented if the second part of the experiment is selected for payment. Evidence of deception - children reporting a different shape with respect to the first that appears – is analyzed with respect to other-regarding preferences and according to the children's ages and genders.

1.3 Results

1.3.1 For 7 to 14 year old children, selfishness diminishes and efficiency concerns increase

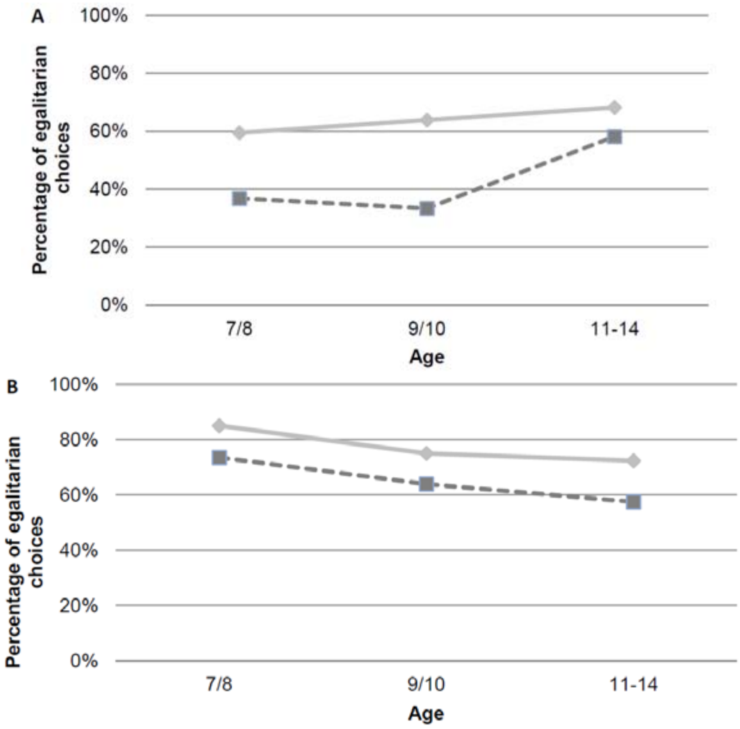
Fig. 1.1 shows the percentage of children choosing the equal sharing (5,5) in the Selfishness, Altruism and Efficiency treatments in the first part of the experiment. Equal sharing is preferred by 48% of the 7-8 year old children in the Selfishness treatment and by 79.73% of them in the Efficiency treatment. In line with previous studies[22-25] we find that as children age, selfishness becomes less prominent and more children exhibit preferences for fairness and efficiency. In the Selfishness treatment, 50% of the 9-10 year old children choose the egalitarian instead of the selfish option, while a higher fraction of participants aged 11 and 14 (63.22%) is prone to making the same choice (p=0.079; χ^2-test, two-sided, N=173). The time path of egalitarian choices is reversed when we consider the Efficiency treatment (see probit regression in Table A3, Appendix): the 11 and 14 year olds are more likely to prefer the (5,7) option compared to the 7-8 year olds (p=0.035; χ^2-test, two-sided, N=168). Indeed, while only 20.27% of younger children choose the efficient alternative, this percentage increases to 30.28% for the 9-10 year olds and to 35.11% for the 11 and 14 year olds. The rate of egalitarian choices differs significantly from an independent random choice in all age groups (binomial test; p=0.005, N=94; p<0.001, N=109, p<0.001, N=74, respectively for 11 and 14, 9-10 and 7-8 year olds). In contrast, in the Altruism treatment 96.40% of the children, independent of age group, prefer equal sharing to the disadvantageous inequality alternative.

Fig. 1.1 | **Frequency of egalitarian choices across age groups.** In the Selfishness treatment, the frequency of egalitarian choices increases with age. In the Efficiency treatment, where the alternative option allows children to augment other's payoff at no cost, the 11-14 year olds are more prone to renounce to share points equally. In the Altruism treatment, most children choose the egalitarian allocation independently of age groups. Error bars illustrate standard errors (SEM).



Girls are more prone than boys to share points equally, independent of the available alternative option (see probit regression in Table A4, Appendix). In the Selfishness treatment, 64.06% of females but only 43.33% of males choose to share points equally (p=0.001; χ^2-test, two-sided, N=248). Similarly, in the Efficiency treatment, 77.04 % of females but only 64.79% of males choose (5,5) (p=0.025; χ^2-test, two-sided, N=277), showing that females are more averse than males to disadvantageous inequality. In the Altruism treatment, 100% of females and 91.84% of males prefer the egalitarian to the altruistic option (p=0.035; *Fisher's exact* test, two-sided, N=111). Some of these gender differences reduce as children age. In the Selfishness treatment, 7-8 year old females more frequently renounce advantageous inequality than males: 59.46% of females and 36.84% of males chose the egalitarian option. This gender gap increases at age 9-10 (p=0.005; χ^2-test, two-sided, N=86) but diminishes strongly at ages 11 and 14 (Fig. 1.2A). Indeed, the percentage of egalitarian choices increases from 33.33% at age 9-10 to 58.14% at ages 11 and 14 for males, while for females it does not significantly change as age increases, going from 63.83% to 68.18%. In contrast, in the Efficiency treatment, the gender gap remains almost constant across ages (Fig. 2B) with efficiency concerns increasing for both males and females: 26.47% of males aged 7-8 and 42.55% at ages 11 and 14 prefer (5,7) to (5,5) while the corresponding percentages for females are 15% and 27.66%.

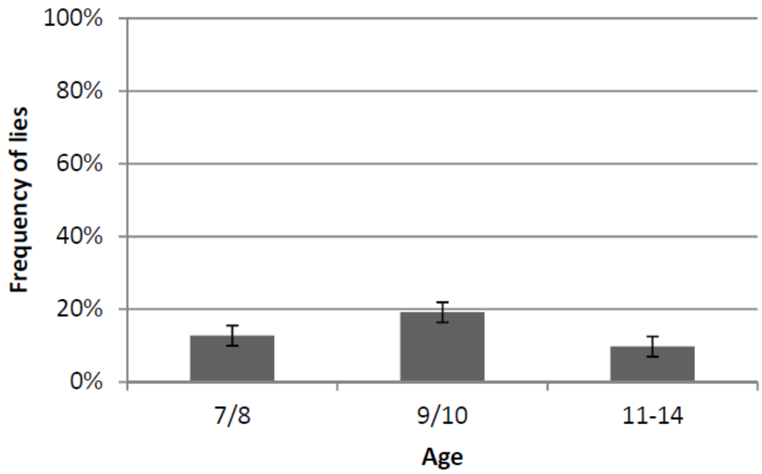
Fig. 1.2 | **Gender differences in egalitarian choices**. The solid lines represent the percentages of egalitarian choices of females; the dashed lines represent the corresponding percentages for males. In the Selfishness treatment (A) the choice is between (5,5) and (7,3) whereas in the Efficiency treatment (B) the choice is between (5,5) and (5,7). (A) shows that boys become less selfish as they age; girls are more likely to share equally at all ages but the gap reduces at older ages. (B) shows that both males and females develop a concern for efficiency as they age; the gender gap is small but persistent.



1.3.2 Deceit follows a hump shaped time path

In the second part of the experiment, 50.08% of the children did not observe their preferred allocation option and thus had an incentive to lie. Among these, only 14.42% lied to obtain their favored outcome. Such low incidence of deceitfulness suggests that most children are lie-averse. The incidence of children's lies depends on their age, social preferences, and gender. Fig. 3 displays the distribution of lies by age group. 12.75% of children at age 7-8 lied to get their preferred option. This is roughly 6 percentage points lower than the corresponding fraction of 9-10 year olds (19.20%) but this difference is not significant (p=0.190; χ^2-test, two-sided, N=227). In contrast, 9.78% of the 11 and 14 year olds lied. Deceit thus follows a hump-shaped path across age groups with 9-10 year olds being significantly more prone to lie than older children (p=0.056; χ^2-test, two-sided, N=217).

Fig. 1.3 | **Frequency of lies across age groups**. The figure shows that the evolution of deceit with respect to age groups is hump-shaped. The 9-10 year old children are more likely to lie than older children. Error bars illustrate standard errors (SEM).

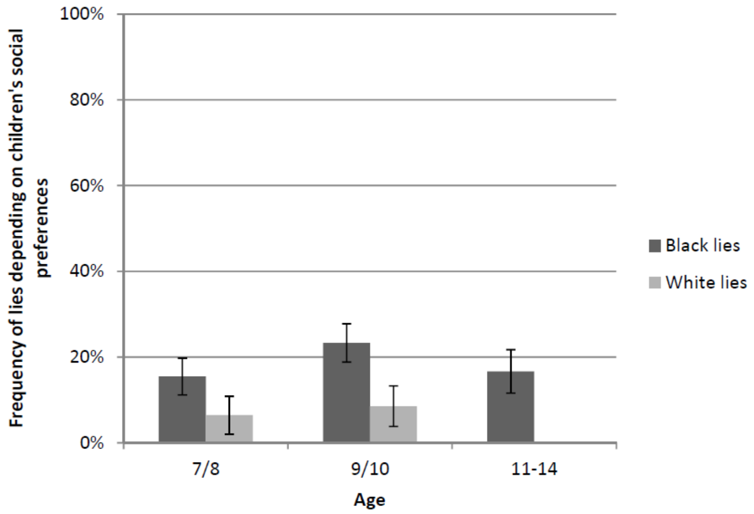


1.3.3 Black lies from selfish children are more frequent than white lies from other-regarding children

To study whether having a self- or an other-regarding preference affects children's willingness to lie, we consider two types of lies with respect to their consequences[15]. *White lies* are defined as lies that benefit the other person; *black lies* are defined as lies that harm the other party. In the Selfishness treatment, a black lie - the child untruthfully reports the (7,3) allocation option instead of the observed one (5,5) - provides a monetary benefit to the liar while reducing his/her partner's payoff. In contrast, a white lie - the child untruthfully reports the (5,5) allocation when observing the (7-3) one - equalizes payoffs by diminishing the liar's payoff and increasing the recipient's payoff. In the Efficiency treatment, a black lie - the child untruthfully reports (5,5) instead of (5,7) - decreases the other's payoff while providing no monetary benefit to the liar. A white lie - untruthfully reporting (5,7) - allows the liar to increase the recipient's payoff at no monetary cost for himself. In the Altruism treatment, a black lie - the child untruthfully reports (5,5) instead of (3,7) - allows the child to increase his/her own payoff at the expense of his/her partner, while a white lie - untruthfully reporting (3,7) - benefits the partner to the liar's detriment.

In the Selfishness treatment, 20% of those who preferred (7,3) in the first part of the experiment but observed the (5-5) option in the second part, told a black lie, while only 4.84% of those who preferred (5,5) to the (7,3) option told a white lie (p=0.014; *Fisher's* Exact test, two-sided, N=127). In the Efficiency treatment, 18.63% of those who preferred (5,5) to (5,7) but did not observe their favored option in the second part, told a black lie while only 5.13% of those who preferred (5,7) lied to benefit their partner although this white lie entailed no personal monetary cost (p=0.062; *Fisher's* Exact test, two-sided, N=141). In the Altruism treatment, no child tells a white lie but 18.75% of those who prefer the (5,5) to the (3,7) option tell a black lie (see Table A1, Appendix). Thus, selfish and envious children are more likely to lie than other-regarding children (see probit regression in Table A6, Appendix). Indeed, children prefer to equalize payoffs by lying to avoid the other's payoff exceeding their own (5,7 or 3,7), while in the Selfishness Treatment, they are prone to renounce equal sharing if lying maximizes their own payoff (7,3).

Fig. 1.4 | **Lying behavior and other-regarding preferences.** Black lies decrease the other's payoff while increasing or holding the liar's payoff constant. White lies increase the other's payoff while reducing or holding the liar's payoff constant. When children have to lie to get their preferred allocation option, they are more likely to tell black lies than white lies. This behavior remains constant across all age groups. Error bars illustrate standard errors (SEM).



Although children's concern for others' payoffs develops with age, we find no statistical evidence of an evolution in white vs. black lie telling behavior with respect to age. Indeed, across all age groups envy and selfishness are more likely to generate lies than other-regarding preferences (see Fig. 4). Among the children who did not observe the shape corresponding to their preferred outcome, 16.67% at ages 11 and 14 tell a black lie (reporting equal sharing in the Efficiency and the Altruism treatments and (7,3) in the Selfishness treatment). In contrast, not one child who, in the first part of the experiment, preferred (5,5) in the Selfishness treatment, (5,7) in the Efficiency treatment, or (3,7) in the Altruism treatment, was prone to tell a white lie when observing the alternative allocation in the second part. The corresponding percentages for 9-10 year old children are 23.33% for black lies and 8.57% for white lies (not significantly different from the older children: *p=0.105*, *Fisher's* Exact test, two-sided, N=73; *p*=0.340, χ^2-test, two-sided, *N*=144, respectively for white and black lies). The youngest children, like the older ones, lie more for selfish or envy reasons (15.49% of black lies vs. 6.45% of white lies).

1.3.4 The gender gap in black lies already exists in childhood

Gender differences in deceitful behavior have been observed in adults, with males being more likely than females to lie in order to secure a monetary payoff[41-42]. Accordingly, across all age levels, 16.77% of male children and 12.10% of female children lie to get their preferred outcome (however, the difference is not significant; *p*=0.246; χ^2-test, two-sided, *N*=319). When we take into account other-regarding preferences, males are more likely than females to tell black lies, and the gender difference in deceit is significant for those children who expressed selfish preferences in the first part of the experiment (*p*=0.097; probit regression in Table A7, Appendix). 23.58% of males but only 14.68% of females lied when preferring an equal sharing in the Efficiency and Altruism treatments or the (7,3) option in the Selfishness treatment (*p*=0.097, χ^2-test, two-sided, *N*=215). In contrast, females tell more white lies than males but not significantly more (3.57% of males and 6.25% of females tell white lies; *p*=0.660; *Fisher's* Exact test, two-sided, *N*=104).

Fig. 1.5 | **Gender differences in black lies**. This figure displays the percentage of children by gender and age group who tell a lie that decreases the other's payoff, among those who did not observe their preferred allocation option in the second part of the experiment. The dashed line is males; the solid line is females. Males are more likely to lie than females, but this difference disappears for older children.

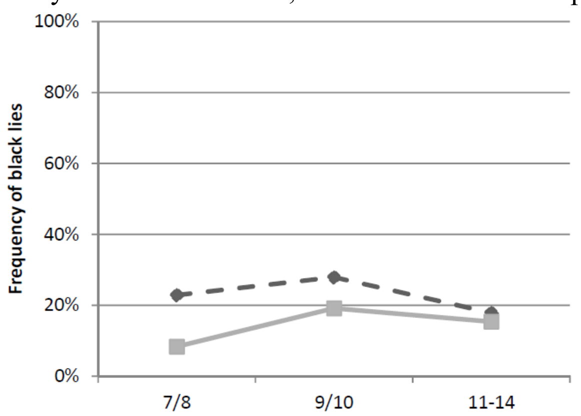


Fig. 5 shows the frequency of black lies by gender with respect to age categories. Boys at age 7-8 seem to be more likely than girls of the same age to tell black lies: 22.86% of boys lied to get more than their partner or to decrease their partner's payoff, whereas only 8.33% of females do so. However, this difference is not significant (p=0.111, *Fisher's* Exact test, two-sided, N=71) and almost disappears in the oldest children, with 17.86% of 11 and 14 year old boys and 15.38% of girls of the same age tell black lies (see Table A2, Appendix).

1.4 Discussion

The development of other-regarding preferences is crucial for supporting cooperation among kin and non-kin individuals in society[3,7,11]. Alongside, honesty is a fundamental dimension of human social life that allows individuals to trust each other and guarantees the integrity of institutions[11-12,44]. Yet, lying is not always considered as an antisocial behavior since some lies can benefit others[15-17]. This study contributes to recent investigation of the role played by age and gender in the development of other-regarding preferences. It identifies how these preferences affect dishonest behavior, especially when children face a moral dilemma such that being generous to others involves lying. Important developmental changes occur in children between the ages of 6 and 14. On entering school, social comparisons and interactions with peers become prominent dimensions of a child's development: through aging, children progressively internalize other-regarding preferences[19-25]. In early adolescence, children become more concerned about efficiency and are more likely to be generous with others even if this implies renouncing an equal share[22-25]. Children progressively base their choices not only on personal gain but also on how their actions will benefit or hurt others. Together with the development of other-regarding preferences, children develop the cognitive ability to lie[37-39]. Although children are taught that lies are morally inappropriate, they learn that deceit, in some context, can be considered morally acceptable, if lying avoid others being harmed.

Overall, however, we observed a lower propensity to lie in children at all ages compared to most studies on cheating behavior in adults[8-10]. This suggests that the norm of honesty is stronger in children and that most view rules as moral absolutes[45]. We acknowledge that conducting the experiment in schools where respect for rules is important could lead to underestimation of children's propensities to lie, since the decision-making environment may influence the individual decision to behave honestly or not[46]. We also did not vary the magnitude of the lie in terms of the monetary consequences for both the liar and the receiver of the lie, which might also explain some of the lie aversion[15,17]. In our experiment, the majority of liars are children who initially expressed selfish preferences and who violated moral rules to satisfy their desire for individual gain or for reasons of envy. The evolution of lying behavior follows an inverted U-shaped time path. Older children who have developed stronger social preferences are less likely to lie than those in middle childhood, even when a lie could benefit the partner at no personal monetary cost. These results suggest that older children internalize the moral norm of honesty more: they are better able than younger children to refrain from a willingness to tell lies for selfish reasons and they do not tell more white lies to implement their social preferences. During early adolescence, children learn to express their other-regarding preferences in ways that do not require lying. In particular, their increasing socialization and the development of internal evaluative reactions, lead them to expect self-disapproval when disregarding their internalized moral standards[38].

References

1. Boehm, C. *Hierarchy in the Forest: The Evolution of Egalitarian Behaviour* (Harvard Univ. Press, Cambridge, 1999).

2. Henrich, J. *et al,.* In search of Homo Economicus: behavioral experiments in 15 small-scale societies. *American Economic Review* **91**, 73-78 (2001).

3. Fehr, E. & Fischbacher, U. The nature of human altruism. *Nature* **425**, 785-791 (2003).

4. Dawes, C. T., Fowler, J. H., Johnson, T., McElreath, R. & Smirnov, O. Egalitarian motives in humans. *Nature* **446**, 794-796 (2007).

5. Tricomi, E., Rangel, A., Camerer, C. F. & O'Doherty, J. P. Neural evidence for inequality-averse social preferences. *Nature* **463**, 1089–1091 (2010).

6. Ostrom, E. *Governing the commons: the evolution of institutions for collective action* (Cambridge Univ. Press, Cambridge, 1990).

7. Fehr, E. & Fischbacher, U. Why social preferences matter – the impact of non-selfish motives on competition, cooperation and incentives. *The Economic Journal* **112**, C1-C33 (2002).

8. López-Pérez, R. & Spiegelman E. Why do people tell the truth? Experimental evidence for pure lie aversion. *Experimental Economics,* in press (available at http://rd.springer.com/article/10.1007/s10683-012-9324-x).

9. Fischbacher, U. & Föllmi-Heusi, F. Lies in disguise – An experimental study on cheating, forthcoming in *Journal of the European Economic Association* (2013).

10. Lundquist, T., Ellingsen, E. & Gribbe, M. Johannesson. The Aversion to Lying. *Journal of Economic Behavior and Organization* **70**, 81-92 (2009).

11. Nyberg, S. The honest society: Stability and policy considerations. *Journal of Public Economics* **64**, 83-99 (1997).

12. Hilbig, B.E. & Zettler, I. Pillars of cooperation: Honesty–Humility, social value orientations, and economic behavior. *Journal of Research in Personality* **43**, 516–519 (2009).

13. Kant, I. *Groundwork of the metaphysics of morals* (See translated by Cambridge University Press, New York 1998).

14. Gino, F. & Pierce, L. Dishonesty in the name of equity. *Psychological Science* **21**, 1153-1160 (2009).

15. Erat, S. & Gneezy, U. White Lies. *Management Science* **58**, 723-733 (2012).

16. Hurkens, S. & Kartik, N. Would I lie to you? On social preferences and lying aversion. *Experimental Economics* **12**, 180-192 (2009).

17. Gneezy, U. Deception: The Role of Consequences. *American Economic Review* **95**, 384-394 (2005).

18. Ariely, D. *The honest truth about dishonesty* (HarperCollins, 2012).

19. Harbaugh, W. T., Krause, K., Liday, S. G. Jr. & Vesterlund, L. *Trust in children*. In Eds. *Trust, reciprocity and gains from association: Interdisciplinary lessons from experimental research* (E. Ostrom, J. Walker, Russel Sage Foundation, New York, 2003).

20. Benenson, J. F., Pascoe, J. & Radmore, N. Children's Altruistic Behavior in the Dictator Game. *Evolution and Human Behavior* **28**, 168-175 (2007).

21. Sutter, M. & Kocher, M. G.. Trust and trustworthiness across different age groups. *Games Econ. Behavior* **59**, 364-382 (2007).

22. Fehr, E., Bernhard, H. & Rockenbach, B. Egalitarianism in young children. *Nature* **454**, 1079-1084 (2008).

23. Almås, I., Cappelen, A. W., Sørensen, E. Ø. & Tungodden, B. Fairness and the Development of Inequality Acceptance. *Science* **328**, 1176-1178 (2010).

24. Fehr, E., Rützler, D. & Sutter, M. The Development of Egalitarianism, Altruism, Spite and Parochialism in Childhood and Adolescence *IZA Discussion Paper 5530* (2011).

25. Martinsson, P., Nordbloma, K., Rützler, D. & Sutter, M. Social preferences during childhood and the role of gender and age. An experiment in Austria and Sweden. *Economics Letters* **110**, 248-251 (2011).

26. Talwar,V. & Crossman, A. From little white lies to filthy liars: The evolution of honesty and deception in young children. *Advances in Child Development and Behavior* **40**, 139-179 (2011).

27. Bowles, S. Endogenous preferences: the cultural consequences of markets and other economic institutions. *Journal of Economic Literature* **36**, 75-111 (1998).

28. Talwar, V. & Lee, K. Little liars: Origins of verbal deception in children. In Ed. *Origins of the social mind: evolutionary and developmental views* (Shōji Itakura, Kazuo Fujita Springer, 2008), pp. 157-178.

29. Bussey, K. Children's categorization and evaluation of different types of lies and truths. *Child Development* **70**, 1338-1347 (1999).

30. Bucciol, A. & Piovesan, M. Luck or cheating? A field experiment on honesty with children. *Journal of Economic Psychology* **32**, 73-78 (2011).

31. Lewis, M., Stanger, C. & Sullivan, M. W. Deception in 3-Year-Olds. *Developmental Psychology* **25**, 3, 439-443 (1989).

32. Chandler, M. J., Fritz, A. S. & Hala, S. M. Small-scale deceit: Deception as a marker of two-, three-, and four-year-olds' early theories of mind. *Child Development* **60**, 1263-1277 (1989).

33. Broomfield, K. A., Robinson, E. J. & Robinson, W. P. Children's understanding about white lies. *British Journal of Developmental Psychology* **20**, 1, 47-65 (2002).

34. Talwar, V. & Lee, K. Emergence of white lie-telling in children between 3 and 7 years of age. *Merrill-Palmer Quarterly* **48**, 160-181 (2002).

35. Hala, S. & Russell, J. Executive Control within Strategic Deception: A Window on Early Cognitive Develoment? *Journal of Experimental Child Psychology* **80**, 112-141 (2001).

36. Talwar, V. & Lee, K. Social and Cognitive Correlates of Children's Lying Behavior. *Child Development* **79**, 4, 866-881 (2008).

37. Talwar, V, Gordon, H. & Lee, K. Lying in the elementary school: Verbal deception and its relation to second-order belief understanding. *Developmental Psychology* **43**, 804-810 (2007).

38. Bussey, K. Lying and truthfulness: Children's definitions, standards, and evaluative reactions. *Child Development* **63**, 129–137 (1992).

39. Xu, F., Bao, X., Fu, G., Talwar, V. & Lee, K. Lying and truth-telling in children: from concept to action. *Child development* **81**, 581-596 (2010).

40. Talwar, V., Murphy, S. M. & Lee, K. White lie-telling in children for politeness purposes. *International Journal of Behavioral Development* ***30**,* 1-11 (2006).

41. Dreber, A. & Johannesson, M. Gender differences in deception. *Economics Letters* **99**, 197-199 (2008).

42. Childs, J. Gender differences in lying. *Economic Letters* **114**, 147-149 (2012).

43. Fischbacher, U. z-Tree: Zurich Toolbox for Ready-made Economic Experiments. *Experimental Economics* **10**, 171-178 (2007).

44. Cialdini, R. B., Petrova, P. K. & Goldstein, N. J. The hidden costs of organizational dishonesty. *MIT Sloan Management Review* **45**, 67-73 (2004).

45. Piaget, J. *The moral judgment of the child* (The Free Press, New York, 1965).

46. Abeler, J., Becker, A. & Falk, A. Truth-telling – A representative assessment *IZA Discussion Paper* 6919 (2012).

Appendix

A.1 Experimental materials

In order to run the experiment we used:

- a computer (the experiment was programmed and conducted with the software z-Tree)
- a card-board box
- pen and reporting sheets
- a coin
- a white and a black bag.

A.2 Experimental procedures

Each session was run in the classroom. Only children whose parents have signed the consent form could participate in the experiment; moreover, children were explicitly asked if they wanted to participate[1]. Each child participated in one of the following treatments: Selfishness, Efficiency, or Altruism treatment. 34 classes in total participated in the experiment, with an average of 19 students per class (S.D.=3.45). Treatments were assigned randomly to the classes.

Each child was randomly paired with an anonymous classmate and made his decisions alone, in a separate room, while the experimenter was waiting outside. None of the children was informed about the identity of his partner.

[1] In total, our sample consists of 637 children. We contacted 742 parents, and 686 gave their consent (92%). The size of the final dataset is reduced further because 2 children voluntary decided not to participate and 42 children were absent the day of the experiment.

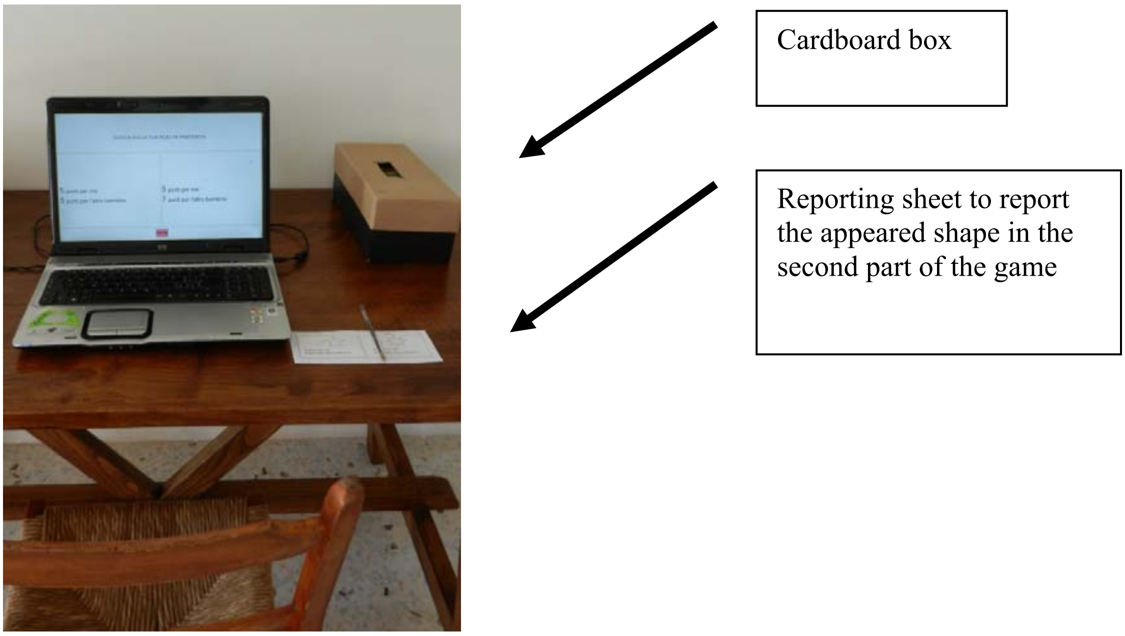


Figure A1. Experimental set up

Figure A1 shows the experimental set up: a computer, the reporting sheet and the cardboard box.

Instructions were explained to all children in the class but questions were asked and answered privately. The children were asked not to talk with their neighbors during the session. After the instructions were read aloud, children made their two decisions one by one, alone, in a separate room. The experiment was divided into two parts: each child participated to both parts but was randomly paid only according to the choice he has made in the first or in the second part of the experiment. Children were told which part was paid only at the end of the experiment. Moreover, they were explained that only the choice made by one of the two members in the pair in the selected part would be randomly selected for payment. In the case there was an odd number of children in the class, one child was randomly selected to play the game without a partner (this was not made common knowledge): one of his two decisions was randomly selected for determining the child's payment. In the payment phase, the earned points were exchanged with prizes (see Figure A5): the higher the number of points earned by the child, the more prizes he would receive.[2]

[2] We implemented a slightly stronger Exchange rate for payments in the 14 years old group. This approach was taken to ensure that the marginal incentives were comparable across ages.

In the first part of the game, each child was asked to privately choose between two allocation options. The two options were shown on a computer screen and the child had to tick his preferred option (see Figure A2). After a decision was taken, the second part of the game began.

In the second part, the child was asked to click on the “appear” button to make a shape appear: either a sun or a star (see Figure A3). The probability that the sun or the star appears is the same and this was made common knowledge. The child was instructed to report the shape displayed on the computer screen on a reporting sheet (see Figure A4). Each shape was associated with one of the two allocation options used in the first part. The allocation option reported by the child was implemented if the second part was randomly selected for payment at the end of the session. We chose to display shapes and not payoff allocations because it has been shown that lying behavior is made easier when the monetary payoff is more "distant" in psychological terms (Ariely, 2012).

Finally, the child had to exit the game by pressing an “Exit” button that closed the windows. Before leaving the room, he had to put the reporting sheet in a cardboard box.

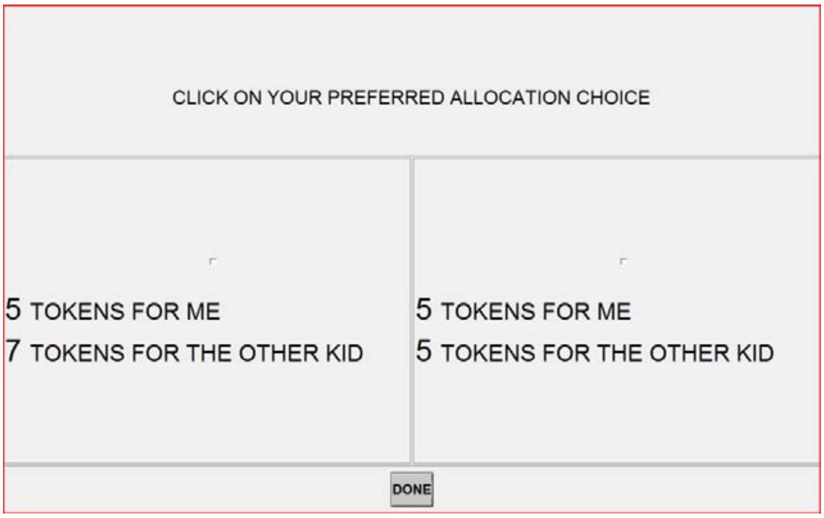
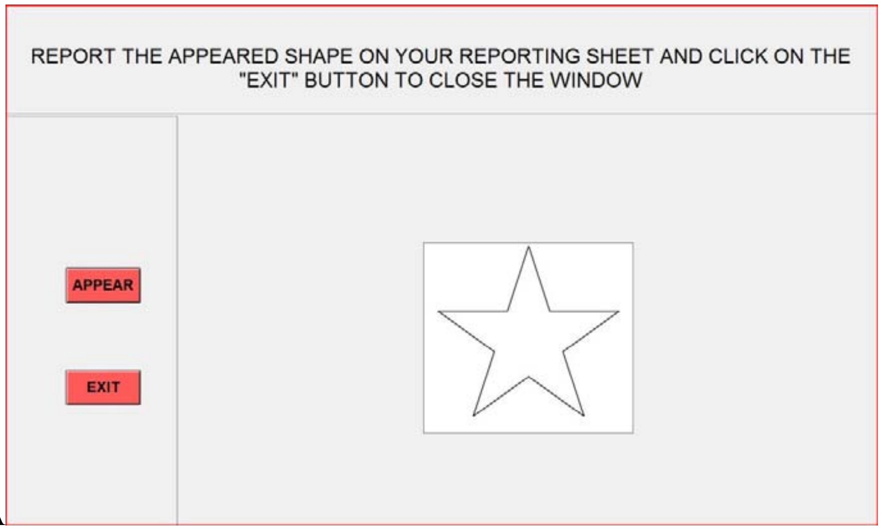
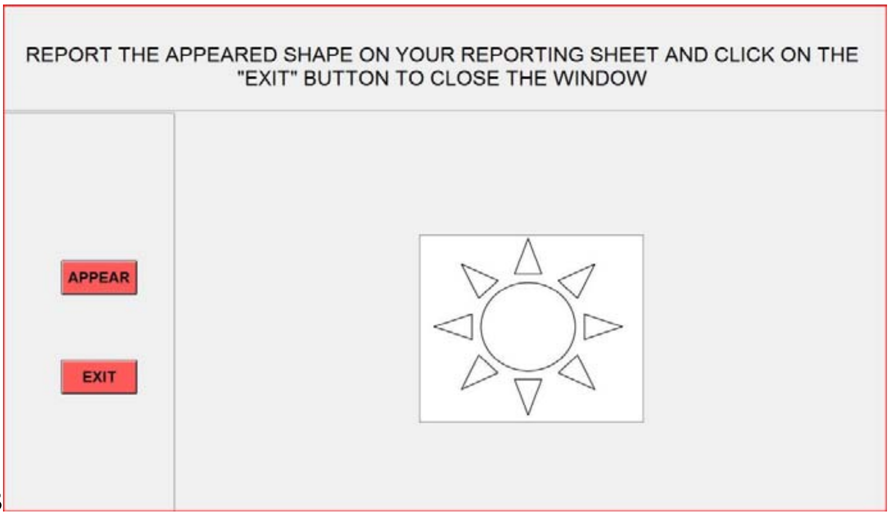


Figure A2. First part of the game (Efficiency treatment)

Figure A2 displays the computer screen for the first part of the game. Children had to choose between two allocation options.



A



B

Figure A3. Second part of the game

Figure A3 displays the computer screen for the second part of the game. Each child had 50% probability to see a star (A) or a sun (B).

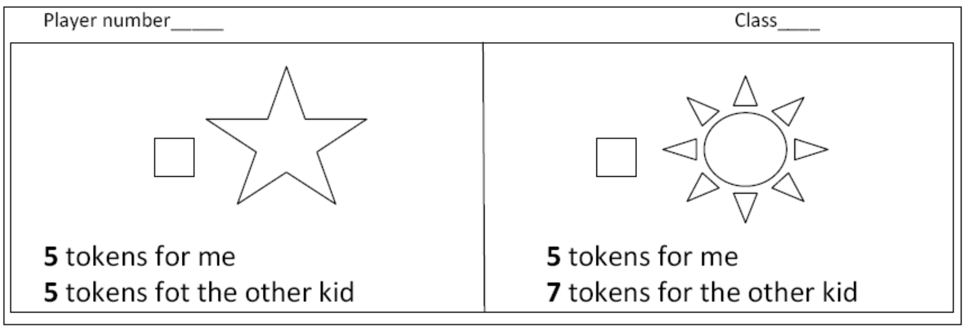


Figure A4. Second part of the game (Efficiency treatment)

Figure A4 displays the reporting sheet for the second part. Each child had to report the appeared shape: depending on the reported shape, one of the two allocation options was implemented.

Figure A5. Prizes

Figure A5 shows the prizes that children could get in exchange for the earned tokens: stickers, colored pens and pencils, “shokky-bandz”.

Instructions for both parts were read in a row: in such a way, children had to go to the other room to play the game only once, reducing the possibility of communicating their choices to their peers during the experiment. This procedure also saved time (the average duration of the game was about 50 minutes). Alternatively, we could have asked children to make their first decision on a sheet of paper in the classroom and then to make their second decision on the computer, one by one, alone, in another room. However, with this alternative procedure, children would have taken their decisions in two different environments: in the classroom, sitting next to their peers, and then in isolation in another room. Our procedure kept the environment constant throughout the game.

Reading the instructions of both parts in a row require that children’s attention remains focused for the entire duration of the instructions. To help children, instructions were briefly explained one more time to each child individually before he entered the decision room (see below). In addition, it was made clear that the shape that appears in the second part was independent of the first decision.

Our experiment was computerized but we think that children's understanding of its functioning was not an issue. Indeed, according to recent research, the average Italian child from age 2 to 11 spends 22 hours a month on the computer and children from age 2 to 5 are more likely to be able to use a computer mouse and play a computer game than tie their own shoelaces (Bricolo, F. *et al.,* 2007; Smith, JR., 2013). Moreover, the experimenter stood nearby the decision room in case the child needed help. No child has encountered any problem with the computer.

A.3 Experimental instructions *(translated from Italian)*

1. Introduction

Hi everybody! How are you? (*children answer*) My name is Valeria, this is Daniela and this is Giulia. First of all, thanks a lot for letting us come to your school today. It is really nice to be here. Today we are going to play a simple game! Only children whose parents have signed the consent form will participate in the game. If your parents have not signed the consent form, please stay quiet. You are free to not participate in the game if you do not want to, just raise your hand and wait.

2. Identification

First of all, you will randomly pick a tag with a random number on it, because it is too difficult for us to remember all your names! Please, attach your number to your shirt. You are not allowed to change your number with others. Please keep this number with you until the end of the game.

Each child picks a number from a bag.

Are you ready? If you do not understand something of this explanation, do not worry: before playing the game I will individually explain to each participant the rules of the game one more time. I will then privately answer all your questions.

3. Explanations of the rules to the class

Let's start by explaining the rules of the game. Please pay attention and be quiet! In this game you will have the possibility to earn some tokens that you can exchange at the end of the experiment for pencils, shokkhy-bandz, etc.

An assistant shows the 'prizes' to children (see Figure A5).

As you can see, there are enough pencils, shokky-bandz, etc. for everybody. Obviously, the more tokens you have, the more things you can get.

This game consists of two parts and you will get the tokens you have earned in just one of these two parts: this part will be randomly selected at the very end of the game. Which part to consider in order to give you the tokens DOES NOT DEPEND on the choices you will take, it is completely random.

Now, I am going to explain you the first part of the game: please pay attention and be quiet. If you have any questions I will answer them privately and individually before playing the game.

This game is played in pairs. You are going to be randomly paired with one of your classmates. You will never know which child is your partner in this game.

Remember that you cannot choose your partner, s/he will be chosen at random.

You will play this game one by one, using a computer. Do not worry, this game is really simple, and if you have any doubts, I will answer all your questions before you play it!

In the first part of the game, you have to make a decision: you have to decide how to divide 10 tokens between yourself and your partner. You will take this decision alone, and your parents, your classmates, your teachers will never know what was your choice, not even at the end of the game. One by one, you will leave the classroom, go to another room and make your decision by clicking on the computer screen.

This is what you will see on the computer:

One assistant shows the poster for the first part of the game (see Figure A2) to the children.

As you can see, there are two possible ways to allocate the tokens between you and the other kid: the option on the left-hand side and the option on the right-hand side. The option on the left-hand side will give to you and to the other kid the same amount of tokens, that is 5, whereas the right-hand side option will give 5/7/3 tokens to you and 7/3 tokens to the other kid. You have to tick the option you prefer by clicking on it. After you have ticked your preferred option, you have to click on the "done" button and wait. Remember that there are no correct or incorrect choices, just choose the option that you prefer.

You will take this decision alone. The other children, the teacher, your friends and your parents will never know what was your choice. Remember that your choice is private, so you are not allowed to speak about it until the very end of the game.

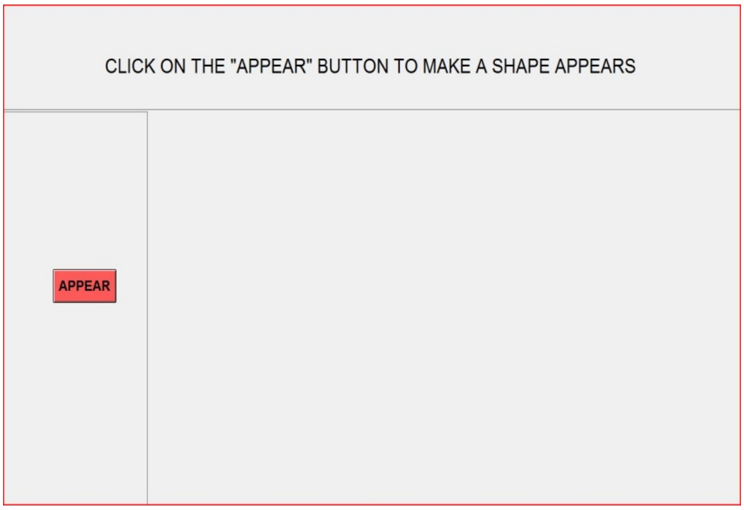
The explanation of the first part of the game is finished…as you have seen, it is a very simple game!

If you have not understood what you have to do, do not worry, before playing the game I am going to individually explain to you the rules of the first part of the game and I will answer all your questions.

Let's start with the second part of the game! Are you ready? **Please remember that you are going to get only the tokens you have earned in only one of the two parts of the game: this part will be randomly selected!**

After you have clicked on your preferred allocation choice and on the "done" button in the first part of the game, the second part of the game begins. In the second part of the game you will firstly see this screen.

An assistant shows the following poster showing the screenshot for the second part of the game



As you can see, this screen asks you to click on the "APPEAR" button in order to make a picture appear. Which picture? A sun or a star! As you click on this "APPEAR" button you will see a sun or a star. In the computer there are as many suns as stars, so you will see one of the two at random. Some of you will see a star, some others will see a sun.

An assistant shows the sun and the star (Figure A3) to the children.

Once you have seen the shape, you have to report which picture you have seen in this reporting sheet.

An assistant shows the reporting sheet (figure A4) to the children

As you can see, this reporting sheet is similar to the computer screen I have just shown you in the first part of the game. The only difference is that you have two pictures: a sun and a star, one for each of the two options.

If you tick that you have seen the star then both you and the other kid will receive 5 tokens, whereas if you tick you have seen the sun then you will receive 5/7/3 tokens and the other kid will receive 7/3 tokens.

After you have ticked one of the two options in the reporting sheet, you have to exit the game by clicking on the "EXIT" button: the "EXIT" button will close the computer screen. Each of

you will tick one of the two options alone and the other children, the teacher, your friends and your parents will never know which option you ticked.
Once you have clicked on the exit button the screen will go grey and the game will be finished!

After that, you have to put the reporting sheet in the cardboard box and then come back to your classroom. As in the first part of the game, you are not allowed to speak about it until the very end of the game.

At the very end of the game, in order to give you the tokens, we will toss this coin in order to select randomly whether we will consider the decisions made in the first part of the game **or** the reporting sheets you put in the cardboard box in the second part of the game.

As you can see, in these two bags – a white and a black bag – we have all your numbers: numbers in the white bag are paired with numbers in the black bag. Since all of you have participated in the game, at the end of the experiment we will select randomly between one of the two bags and will distribute the tokens according to the choices these children made.

Remember, however, that you will never know the identity of your partner.

If you do not understand what you have to do, don’t worry, before playing the game I am going to individually re-explain to you the rules of the first and the second parts of the game and I will answer all your questions.

Let’s start with the game! You can come into the room next door, one by one, and play the game with us. Giulia will call out numbers and when you hear your number it's your turn to play. Giulia will take you to me, where you will play the games. When you are finished the games, you can go back to the classroom and carry on drawing. It’s really important to us that you do not talk about the game until all the kids have played. We really want you to follow this rule!

4. Children play the game

One assistant calls out the next number and fills in name, surname and number of the child on a table. Each odd numbered child is matched with an even numbered child.

Individual explanation of the rules of the game

Hi! Now I shall explain to you the rules of the game. If you have any questions, please ask me, I shall be happy to answer them all!

In this game you will be paired up with one of the children in your class. You will never know the identity of the other kid, and the other kid will never know your identity, not even at the end of the game. You have to decide how to divide some tokens between yourself and the other kid.

In the first part of the game, you have to choose between two allocation options and you have to click on the one that you prefer. If you tick the left-hand side option, then both you and the other kid will receive 5 tokens; if you tick the right-hand side option then you will receive 7/5/3 tokens whereas the other kid will receive 3/7 tokens. Your friends, your teacher, your classmates and your parents will never know what you chose.

Is it clear? Do you have questions?

After you have clicked on your preferred choice you have to click on the “Done” button. The first part of the game is finished.

In the second part of the game, you have to click on the “Appear” button to make a sun or a star appear. After you have seen the picture, you have to report it on this reporting sheet.

The child is given a reporting sheet, with his/her own number on it.

If you report the sun, then you and the other kid will both receive 5 tokens; if you report the star then you will receive 7/5/3 tokens whereas the other kid will receive 3/7 tokens. After you have reported the appeared shape in the reporting sheet, you have to click on the EXIT button to exit the game and close the video screen. Finally, you have to put your reporting sheet in the cardboard box.

Your friends, your teacher, your classmates and your parents will never know which option you ticked.

Remember that you will receive the tokens according to the decisions you make either in the first **or** in the second part of the game, and this will be random.

Is it clear? Do you have questions?

5. Distribution of the 'prizes', after all children have played the game

After all kids have played the game, one assistant tosses a coin to determine whether children will be paid according to the choices made in the first or in the second part of the game. The coin is tossed again to determine whether children will be paid according to the choices made by the children whose number is contained in the white bag or by the children whose number is contained in the black bag.

Now that all the children have played the game, we will tell each of you how many tokens you have won and we will exchange them for what you want, according to this rule:

The following exchange rule is written on the blackboard.

> 3 tokens: you can get 1 colored pen/pencil or 1 packet of stickers or 5 shokky-bandz.
>
> 5 tokens: you can get 2 colored pens/pencils or 2 packets of stickers or 10 shokky-bandz. You can also mix the prizes and get 1 pen/pencil and 1 packet of stickers or 1 packet of stickers and 5 shokky-bandz, etc.
>
> 7 tokens: you can get 3 colored pens/pencils or 3 packets of stickers or 1 packet of shokky-bandz. You can also mix the prizes and get 1 colored pen/pencil, 5 shokky-bandz and 1 packet of stickers or 2 packets of stickers and 5 shokky-bandz, etc.

When I call your number please come here: I will tell you how many tokens you have won so that you can choose the prize you want.

Once you get the prize, please put it in your schoolbag and keep quiet.

6. Individual distribution of the prize

Hi, you have won 3/5/7 tokens! You can get these items, what do you want? Do you like the prize?

7. End of the experiment

Thank you very much for participating in this game! We really enjoy our staying! Now you will continue your lesson with Professor xxx. Please, remember not to talk with the others about the game! Bye bye!

A.4 Summary statistics

The following tables contain the summary statistics on the lying behavior of participants. Lying behavior is analyzed depending on their other-regarding preferences, as revealed in the first part of the game of each treatment, by age group in Table A1 and by gender in Table A2.

Table A1. Relative frequency of lies of children who did not observe in Part 2 their preferred allocation option as revealed in Part 1, by age groups, type of lie, and treatment.

Relative frequency in %	Black lies			White lies		
Treatment	Self.	Effic.	Altruism	Self.	Effic.	Altruism
Actually observed option	(5,5)	(5,7)	(3,7)	(7,3)	(5,5)	(5,5)
Reported option in Part 2	(7,3)	(5,5)	(5,5)	(5,5)	(5,7)	(3,7)
7/8 years old (a)	12.50	15.15	21.43	9.52	0	0
9/10 years old (b)	28.00	23.08	19.23	5.00	14.29	0
11-14 years old (c)	18.75	16.67	12.50	0	0	0
Fisher's exact test (a-b)	0.289	0.561	1	1	0.502	/
Fisher's exact test (a-c)	0.668	1	1	0.488	/	/
Fisher's exact test (b-c)	0.712	0.552	1	0.488	0.209	/
Total	20.00	18.63	18.75	4.84	5.13	0

Table A2. Relative frequency of lies of children who did not observe in Part 2 their preferred allocation option as revealed in Part 1, by gender, type of lie, and treatment.

Relative frequency in %	Black lies			White lies		
Treatment	Self.	Effic.	Altruism	Self.	Effic.	Altruism
Actually observed option	(5,5)	(5,7)	(3,7)	(7,3)	(5,5)	(5,5)
Reported option in Part 2	(7,3)	(5,5)	(5,5)	(5,5)	(5,7)	(3,7)
Males	21.43	23.91	27.78	3.57	4.00	0
Females	17.39	14.29	13.33	5.88	7.14	0
Fisher's exact test (χ^2-test*)	0.758	0.214*	0.265	1	1	/
Total	20.00	18.63	18.75	4.84	5.13	0

In both Tables A1 and A2, the relative frequency of lies is reported with respect to the children who had an incentive to lie, i.e. those who actually observed in the second Part of the game the shape that does not correspond to their preferred allocation option, as revealed in the first Part. It is important to note that the probability of cheating by those children who had an incentive to lie (14.42%) is significantly higher than the probability to misreport the observed shape by those children who did observe the shape corresponding to their preferred option (3.42%) (p<0.001, normal approximation two sample test of equality of independent proportions, two sided, N=637) .

In the Selfishness treatment, a **black lie** decreases the other's payoff, while increasing the monetary outcome of the liar. The child reports the (7-3) option while the appeared shape corresponded to the (5,5) allocation. In the Efficiency treatment, children tell a black lie when they report (5,5) whereas they observed the shape corresponding to the (5-7) option. In the Altruism treatment, they tell a black lie when they report (5-5) whereas they observed the shape corresponding to the (3-7) option.

A **white lie** provides a benefit to the other part. Children tell a white lie when they report the (5-5) option instead of the observed (7-3) option in the Selfishness treatment. In the Efficiency treatment, the children tell a white lie when they report (5,7) whereas the shape corresponded to the (5,5) option. In the Altruism treatment, a white lie (untruly reporting (3-7) instead of (5-5)) is very costly for the child since he has both to renounce to equally distribute points and to incur in a disadvantageous inequality.

A.5 Regression analysis

The following table contain summary of Probit regressions used in the paper.

<u>Notation</u>

Throughout the tests we use "(5-5) choice" and "Egalitarian choice" as synonyms.

The *Egalitarian choice* variable is a dummy variable that takes value 1 if (5-5) is chosen, and 0 otherwise.

The *Lie* variable is a dummy variable that takes value 1 if the child has misreported the observed allocation option when having an incentive to misreport (i.e. among the subjects who actually observed the shape that did not correspond to their preferred allocation option), and 0 otherwise.

The *Selfishness treatment* variable is a dummy variable that takes value 1 if the child participates in the Selfishness treatment, and 0 if he participates in the Efficiency treatment. The Altruism treatment is not considered in the analysis of the choices taken by children in the first part of the game (Tables A3 to A5) since in our article we have provided evidence that the observed behavior in the Altruism treatment did not considerably differ with respect to children's age and gender, our main independent variables of interest. It is thus more informative to estimate the Probit regression focusing on the Selfishness and Efficiency Treatments.

The *Female* variable is a dummy variable that takes value 1 if the decision-maker is a female, and 0 otherwise.

The *Young age*, *Medium age* and *Older age* variables refer to the three age groups, respectively 7/8, 9/10 and 11-14 years old. These variables are dummy variables that take value 1 if the child belongs to the defined age group, and 0 otherwise.

The *Other-regarding choice* is a dummy variable that takes value 1 if the decision-maker expresses other-regarding preferences in the first part of the game, and 0 if the child prefers the self-regarding option, i.e. (7,3) in the Selfishness Treatment or the equal sharing in the Altruism and Efficiency Treatments.

Tables with regression analysis

<u>*Egalitarian choices and age*</u>

We have estimated a Probit model in which the dependent variable is the Egalitarian choice. The independent variables are *Young age, Medium age, Selfishness treatment*, and interaction terms between *Young age*, *Medium age* and *Selfishness treatment* (N=525). The following Table reports the marginal effects and their level of significance.

Table A3. Regression analysis: egalitarian choice and age

Young age	*dF/dx* = 0.162, *p* = 0.034, *z* = 2.12
Medium age	*dF/dx* = 0.040, *p* = 0.555, *z* = 0.59
Selfishness treatment	*dF/dx* = - 0.017, *p* = 0.814, *z* = - 0,23
Selfishness treatment * Young age	*dF/dx* = - 0.324, *p* = 0.004, *z* = - 2.87
Selfishness treatment * Medium age	*dF/dx* = - 0.173, *p* = 0.093, *z* = - 1.68
Log likelihood: -334.639	

Note that the omitted variables in this regression are *Older age* and *Efficiency treatment*. Thus, the marginal effects of the *Young age* and *Medium age* variables are measured relative to the oldest age group in the Efficiency treatment (i.e. the 11-14 years old children). Therefore, the *Young age* variable measures the (negative) age trend in choosing the egalitarian choice in the Efficient Treatment. Finally, the interaction term Selfishness treatment * Young age measures the difference in the probability of choosing the egalitarian choice between the Selfishness and the Efficiency treatments in the youngest group versus the difference in the oldest children: a negative coefficient means that the difference between the Selfishness treatment and the Efficiency treatment declines with age, *i.e.* as children become older, they choose less often the (5-5) option in the Efficiency treatment and more often the (5-5) option in the Selfishness treatment.

<u>*Egalitarian choices and gender*</u>

We have estimated a Probit model in which the dependent variable is the Egalitarian choice. The independent variables are *Female, Selfishness treatment*, and an interaction term between *Selfishness treatment* and *Female* (N=525). The following Table reports the marginal effects and their level of significance.

Table A4. Regression analysis: egalitarian choice and gender

Female	*dF/dx* = 0.142, *p* = 0.018, *z* = 2.36
Selfishness treatment	*dF/dx* = - 0.198, *p* = 0.001, *z* = - 3.36
Selfishness treatment * Female	*dF/dx* = 0.055, *p* = 0.514, *z* = 0.65
Log likelihood: -331.157	

Note that the interaction term Selfishness treatment * Female is not significant. The impact of children's gender on the probability to choose the egalitarian option is the same regardless of the treatment, *i.e.* females choose (5,5) more often than males both in the Selfishness and in the Efficiency Treatment.

Egalitarian choices, gender and age

We have estimated a Probit model in which the dependent variable is the Egalitarian choice. The independent variables are *Young age, Medium age, Female, Selfishness treatment*, and interaction terms between *Young age*, *Medium age* and *Selfishness treatment* (N=525). The following Table reports the marginal effects and their level of significance.

Table A5. Regression analysis: egalitarian choice, age and gender

Young age	*dF/dx* = 0.159, *p* = 0.039, *z* = 2.06
Medium age	*dF/dx* = 0.051, *p* = 0.459, *z* = 0.74
Female	*dF/dx* = 0.169, *p* = 0.000, *z* = 3.96
Selfishness treatment	*dF/dx* = - 0.018, *p* = 0.795, *z* = - 0,26
Selfishness treatment * Young age	*dF/dx* = - 0.322, *p* = 0.005, *z* = - 2.83
Selfishness treatment * Medium age	*dF/dx* = - 0.194, *p* = 0.062, *z* = - 1.86
Log likelihood: -326.714	

Lying decision and children's preferences

We have estimated a Probit model in which the dependent variable is (the decision to) Lie. The independent variable is *Other-regarding choice* (N=318). The following Table reports the marginal effect and its level of significance.

Table A6. Regression analysis: Decision to lie and social preferences

Other-regarding choice	*dF/dx* = - 0.143, *p* = 0.001, *z* = -3.41
Log likelihood: -124.595	

Table A6 provides evidence that the lying behavior of children differs with respect to their distributive preferences. The negative sign of the *Other-regarding choice* variable indicates that children who expressed other-regarding preferences in the first Part of the game are significantly less likely to cheat to obtain their preferred allocation option than the self-regarding children.

Lying decision, children's preferences and gender

We have estimated a Probit model in which the dependent variable is the decision to lie. The independent variables are *Other-regarding choice, Female* and an interaction term between *Other-regarding choice* and *Female.* (N=318). The marginal effect of the *Female* variable is measured relative to the male group when having self-regarding preferences. The following Table reports the marginal effects and their level of significance.

Table A7. Regression analysis: Decision to lie, social preferences and gender

Other-regarding choice	*dF/dx* = - 0.184, *p* = 0.002, *z* = -3.16
Female	*dF/dx* = - 0.068, *p* = 0.097, *z* = -1.66
Female * Other-regarding choice	*dF/dx* = 0.154, *p* = 0.201, *z* = 1.28
Log likelihood: -123.216	

In Table A7, the *Female* variable indicates that females have a lower propensity to tell a lie with respect to boys when being self-regarding.

Lying decision, children's preferences, gender and age

We have estimated a Probit model in which the dependent variable is the decision to lie. The independent variables are *Other-regarding choice, Female*, *Young age* and *Medium age* (N=318). We measure the marginal effects of the *Young age* and *Medium age* variables with respect to the oldest age group (omitted group): the *Middle age* variable measures the (slightly) higher propensity to lie of the children aged eleven and fourteen compared with the nine-ten years old ones. The following Table reports the marginal effects and their level of significance.

Table A8. Regression analysis: Decision to lie, social preferences, age and gender

Other-regarding choice	*dF/dx* = - 0.179, *p* = 0.002, *z* = -3.10
Young age	*dF/dx* = 0.024, *p* = 0.642, *z* = 0.47
Medium age	*dF/dx* = 0.084, *p* = 0.090, *z* = 1.72
Female	*dF/dx* = - 0.069, *p* = 0.086, *z* = -1.72
Female * Other-regarding choice	*dF/dx* = 0.156, *p* = 0.196, *z* = 1.29
Log likelihood: -121.486	

References of the Appendix

1 D. Ariely. *The (Honest) Truth About Dishonesty* (HarpercCollins, 2012).

2 F. Bricolo, D. A. Gentile, R. L. Smelser, G. Serpelloni. Use of the Computer and Internet among Italian families: First national study. *Cyberpsychology & Behavior* **10**, 789-798 (2007).

3 J. R. Smith. Kids are learning computer skills before life skills. Retrieved January 1, 2013, from *http://www.avg.com*

Chapter 2

Workers' Selection, Effort and Conflict of interests

with Natalia Montinari and Antonio Nicolò

2.1 Introduction

One of the main scopes of compensation schemes is to align the interests of the members within an organization. In standard models, alignment of interests between principal and agents is obtained by correlating agents' wages to verifiable outcomes, possibly reducing the efficiency of the job relation due to distortions in risk sharing. However, if agents exhibit reciprocal concerns, the principal can reduce their opportunistic behavior without negatively affecting efficiency, by offering them a wage higher than the incentive compatible one (gift-exchange hypothesis). The fair action by the principal will indeed be reciprocated by a fair action by the agents. It follows that reciprocity concerns are usually considered as desirable features in organizations, reducing the costs of aligning interests between its members.

In hierarchical organizations, however, conflict of interest between its members may emerge at multiple levels: in such a situation, we show that reciprocity can exacerbate, rather than alleviate, the negative effects of members' misalignment of objectives. In particular, reciprocity can induce an "exchange of favors" between some of the members of the organization, decreasing its overall efficiency: one member's intentional actions aimed at illegitimately (i.e. not taking into account the interests of the organization) increasing the welfare of another member will be reciprocated by the latter one, even if this occurs at the expenses of the organization. In this paper, we provide evidence of the emergence of a dark side of reciprocity, which negatively affects organization's profits. In such a situation, the design of the optimal compensation scheme has to conveniently take it into account, in order to (re)align members' interests at all levels of the organization.

We design an experiment that renders a three-level hierarchical organization, formed by one principal, one agent, and two candidates for a job[1]. The agent has to select one worker between the two candidates, who, once hired, chooses a level of non-verifiable effort. The worker can exert his effort either just in activity X, which is beneficial for the whole organization (i.e. both the principal and the agent (who get respectively the 85% and the 15% of the produced value)), or also in activity Y, which inefficiently only provides private benefits to the agent. Importantly, the two candidates differ in their abilities: for each level of effort exerted in activity X, the high ability candidate is more productive than the low ability one, generating higher profits for the organization; however, if exerting effort in activity Y, the two candidates are not different in terms of productivity.

Even if it is public information that candidates have different abilities, only the agent is able to distinguish among them: in particular, the principal cannot observe which one of the two candidates has the higher ability and only receives noisy information regarding the exerted effort by the hired one. We implemented a between subjects design and consider three different treatments: *Baseline, Selection* and *Profit Sharing*. In all treatments, once hired, the worker chooses how much effort to exert in activities X and Y.

In the *Baseline treatment*, the principal orders the agent which candidate to hire: the high or the low ability one. The agent can only obey to the principal and, thus, cannot take any decision (the set of available actions to the agent once the principal has moved is always a singleton).

In the *Selection treatment*, the principal privately suggests to the agent which candidate to hire, but the agent is free to follow the suggestion or not. The fact that the agent takes the hiring decision is made salient by allowing her to send a private message to the selected worker, suggesting how much effort to exert in each activity X and Y.

These two treatments allow us to investigate whether the introduction of a powerful hierarchical level in the organization affects i) the type of candidates who is hired; ii) the level of effort exerted by the hired worker both in activity X and Y. It is important to note that when

[1] We use the female pronoun for the agent and the male pronoun for the principal and for the workers.

the agent thinks that the workers' reciprocal concerns are absent, she will always prefer the high ability to the low ability one, since the former is more productive when exerting the minimum enforceable effort in activity X. If this is the case, we thus should not observe any difference in the hiring choice and in the effort exertion between the *Baseline* and the *Selection* treatment. If on the contrary, the agent tries to take personal advantages from the workers' reciprocal concerns, then we should observe an higher fraction of low ability agents hired in the Selection treatment compared to the Baseline treatment.

Our main findings are twofold. First, in the *Selection treatment*, we observe a substantial and significant increase in the number of low ability candidates hired. Second, we find that such a hiring distortion with respect to the *Baseline treatment* is driven by low ability workers exerting more effort in activity Y than the high ability ones. Agents do strategically exploit the reciprocal concern of low ability workers, who feel less entitled to get the job and are thus more likely to exert effort in activity Y. In particular, in our setting, a compensation scheme that assigns to the agent 15% of the payoff generated in activity X is not high enough to deter the latter to profit of her powerful position at the disadvantage of the organization.

A costly way to avoid the agent to take illegitimate advantage of the reciprocal concerns of the (low ability) workers is to perfectly align her interests with the principal's ones, by increasing the size of the profit sharing to such a level so that she also prefers the worker's effort to be devoted to activity X. As a consequence, both the distortion in the selection process and in the effort provision should disappear. However, we propose a much less costly compensation scheme that helps in reducing the negative effects of conflicts of interests in hierarchical organization: in the *Profit Sharing treatment*, we replicate the design implemented in the *Selection treatment*, with the only difference that the 15% of the monetary payoff generated by the effort exerted in activity X is now equally shared between the agent and the hired worker (7.5% each). We show that this compensation scheme induces both workers to exert a higher level of effort in activity X, while decreasing their effort in activity Y, with respect to the *Selection treatment*. As a consequence, the profit of the organization is restored to the same level as in the *Baseline treatment*, preventing the agent to illegitimately take advantage of her powerful position.

The reminder of this paper is organized as follows. Section 2 sets our paper in the related literature. Section 3 sets out the design and procedures of the experiment. The presentation of our findings is provided in Section 4. Finally, Section 5 discusses our results and concludes.

2.2 Related literature

The gift-exchange hypothesis, first formulated by Akerlof in his seminal paper in 1982 (see also Gintis, 1976; Akerlof and Yellen, 1988, 1990), states that principals offering wages higher than the incentive-compatible ones could be overcompensated for that, with grateful workers exerting more effort than the minimum enforceable level. This hypothesis has received wide support in laboratory (e.g., Fehr et al., 1993, 1997, 1998; Fehr and Gatcher, 1998; Fehr and Falk, 1999; Charness, 2004; Hannan et al., 2002; Brown et al., 2004; Maximiano et al., 2013) and field experiments (e.g., Gneezy and List, 2006; Bellemare and Shearer, 2009; Henning-Schmidt et al., 2010; Kube et al., 2012). This evidence is based on the notion of reciprocity: agents obey to this social norm when perceiving that the offered wage is higher than the (perceived) incentive compatible one. Interestingly, Charness *et al.* (2012) show that delegating the wage choice to employees, rather than just offering them a higher wage, represents a Pareto improvement in the Gift-exchange paradigm. However, whether the "gift exchange" mechanism is beneficial to organizations actually depends on circumstances. Indeed, if reciprocity may positively affect organizations, as in the above described situation, it can also damage them, when the same principle is used by its members as an enforcement device to reach personal illegitimate benefits (Jacquement, 2012). In our study, we are analyzing whether, in a three-level hierarchical organization, such a mechanism will cause workers to inefficiently reciprocate agents who hired them, instead of addressing their effort in favor of the organization they are enrolled in.

When considering the working environment, the hiring process represents a relevant dimension of the principal-agent relationship. In particular, another important thread of the literature is focused on the efficiency loss due to distortions in the selection process: organization's overall performance is usually negatively affected when candidates' evaluation in the hiring process is not based on their ability (Kramarz and Thesmar, 2007; Levine *et al.,* 2010). Managers may indeed favor people according to their personal preferences when objective evaluations of workers' performance is not available (Prendergast and Topel, 1996), or may favor employees who engage in ingratiatory behavior to conform their ideas on managers' opinions, regardless of their objective ability (Robin et al., 2012). In a field experiment, Bandiera *et al*. (2009) show that managers are more likely to hire people according to their social connections when they are paid a fixed wage than when they are paid according to bonuses based on the average productivity of the managed workers.

While all these authors analyze the distortion in the hiring process as a result of the presence of managers' personal preferences towards workers, our first aim is to analyze whether such a distortion is effective even when there are not any kind of social ties among the organizational members. In our experiment, in particular, conflict of interests within the hierarchical organization is a major source in explaining the inefficient employment allocation of workers.

Tirole (1986) and Laffont and Martimort (1997) have expanded the analysis of the traditional principal-agent paradigm to hierarchical organizations, where the principal has to delegate the decision making authority to managers who may, as a result of asymmetric information and costly monitoring, collude with the agents they should supervise. In particular, previous studies analyze how to reduce the negative effects of internal corruption by manipulating the structure of monitoring hierarchies (Mehmet, 1996) and suggest to reduce effort incentives for employees while increasing managers compensations (Thile, 2013). Chang and Lai (2002) investigate the role of social norms in affecting the corruptive behavior of supervisors, showing that when the latter are already plagued with corruption, then paying them more than workers limit worker's slack. Our study contributes to this stream of research providing evidence that, in the presence of conflict of interests between delegated managers and principals, illegitimate behaviors emerges even when managers' compensations scheme is relatively high and related to worker's performance.

In our experiment the agent acts as a hiring expert who has more information than the principal about the quality of the candidates. The negative effect of conflict of interests in professional advisor-client relationships has been analyzed in the medical practice (Dana and Loewenstein, 2003; Loewenstein *et al.*, 2012), in the auditing environment (Cain *et al.*, 2005; Koch and Schmidt, 2010), in the real estate market (Levitt and Syverson, 2008) and in the general framework of experts (Loewenstein *et al.*, 2011; Norton and Isaac, 2012).

Finally, we propose to share a little part of the profit with the workers in order to overcome the hiring and effort distortions caused by the conflict of interests between different levels of the hierarchical organization. Rewards systems do not only represent a monetary incentive to workers but also affect their values, beliefs and attitudes regarding the corporate culture, ultimately affecting their productivity (Kerr and Slocum, 1987).

2.3 Experimental design

In order to test how the presence of a conflict of interest within the organization affects the selection process and the performance of the hired workers, we implement a variation of the design used by Montinari *et al.* (2012), adding one hierarchical level within the organization (i.e. an agent who has to select the worker to be hired) and allowing for inefficient forms of reciprocity, as a result of the presence of conflict of interest between the principal and the agent.

The main game develops as follow. At the beginning of each session, each subject is matched with other three participants to form a group of four, that we call a firm. We will refer to them as Principal (P), Agent (A), Worker L (L) and Worker H (H). First, the principal has the possibility to send a suggestion to the agent, about which worker to hire. After being informed about the suggestion, the agent selects worker L or worker H. The selected worker has then to exert a costly effort in two activities, X and Y, for a fixed compensation. Workers L and H differ only with respect to their productivity, with worker H being more productive than worker L, when exerting the same level of effort in activity X. Only the agent is able to correctly distinguish workers' ability. On the other hand, when plugging away at activity Y, workers have no differences in their ex ante ability. Once selected, the worker has thus to choose a costly effort level in activity X, which produces profits for the entire organization (both the agent and the principal), and a costly effort level in activity Y, which exclusively benefits the agent. However, after selecting one of the two workers, the agent has the opportunity to a priori refuse the value produced in activity Y, if any, thus devoting it to the principal. The cost of effort is increasing in the total effort exerted and does not depend on whether it is exerted in activity X or Y.

We run 3 treatments: *Baseline, Selection* and *Profit Sharing* treatment. In the *Baseline treatment*, the Principal's suggestion whether to hire worker L or H is binding for the agent so that, by design, there is no conflict of interest in this treatment. In the *Selection* and *Profit Sharing treatments*, the agent is free to follow the suggestion of the principal or not. Moreover, in the *Selection* and *Profit Sharing treatments*, the agent has the opportunity to communicate to the selected worker a (non-binding) level of effort she would like him to implement in activity X and Y.

In the *Selection treatment*, it is made clear that the agent is hired by the principal in order to select one worker to work in the organization and should thus act on his behalf, by persecuting the best interest for the organization. The agent, when expecting both workers to exert the minimum level effort (i.e. 1), should then hire worker H, which assures a higher profit for the organization. However, the agent may decide to persecute her personal interest, while disregarding acting on the best interest for the organization. In particular, in such a situation, she may prefer to choose the low ability worker when expecting him to exert more effort in activity Y than the high ability one. Indeed, since the former is less entitled to get the job, he may be more likely to inefficiently reciprocate the agent's choice. Still, as stated by Montinari *et al.* (2012), it is possible that the agent hires the low ability worker because she expects him to work harder for the Principal. In order to distinguish between the illegitimate behavior of the agent and the latter one, we also give the agent the possibility to ex-ante refuse the value eventually produced by the selected worker in activity Y. Moreover, we reasonably think that the agent and the principal hold the same beliefs about workers' effort choices in both activities, so that we expect no differences in their hiring decisions, when they both want to persecute the best interests for the organization. Finally, we also give the agent the opportunity to privately suggest to the hired worker a level of effort to exert both in activity X and in activity Y. Agents may suggest workers to work both for the organization, in activity X, and for themselves, in activity Y, in order to maintain a positive self image; in such a way they behave dishonestly enough to gain but honestly enough to convince themselves of their integrity (Mazar *et al.*, 2008).

The *Profit Sharing treatment* is identical to the *Selection* one with the only difference that the effort exerted in activity X does not only benefits the principal and the agent, but also the worker. When comparing the *Profit Sharing treatment* to the *Selection* one, the conflict of interest between the principal and the agent is still present, but an incentive to act in the interest of the organization is provided to the worker. Interestingly, the cost of such an incentive scheme for the principal is the same as in the *Selection treatment*, where the agent received the 15% of the value produced in activity X: in the *Profit Sharing treatment,* indeed, both the agent and the worker receive the 7.5% of the value produced in activity X. The magnitude of the incentive is very small and it does not affect the optimal effort choice by payoff maximizing workers, equals to the minimum enforceable effort level.

Our experiment consists of two parts. In the first part, participants play the game described above as one shot. In part two, they play the same game for 15 periods, maintaining the same role of part 1 but under a stranger random matching protocol[32]. Participants were informed that the experiment was composed of two parts, but they only received instructions about part two after having completed part one[3]. We can thus analyze whether the above described dark side of reciprocity emerges both in one-shot interactions and as a consequence of players' learning during repetitions, when strategic motives are not at stake.

2.3.1 Payoffs

In the experiment, labor contracts do not contain any explicit incentives to make the worker exert a costly effort. In the *Baseline and Selection treatments*, the worker, once hired, receives a fixed wage w_w equals to 50 ECUs[4], and incurs in a cost when exerting effort in activity X or Y. Differently, in the *Profit Sharing treatment*, the worker also gets the 7.5% of the value he produced in activity X. The chosen effort level in activity X, e_x, has to be an integer number between $\underline{e}_x = 1$ and $\bar{e}_x = 10$, whereas the effort level devoted to activity Y, e_y, has to be an integer number between 0 and 5. The total exerted effort $e_x + e_y$ cannot exceed level 10.

Note that if the worker chooses the minimum enforceable effort level (i.e. 1), this effort is necessarily exerted in favor of activity X. Therefore, workers who aim at maximizing their payoff, will only exert a level of effort equals to 1 in activity X (this is true also for the profit sharing treatment, since the magnitude of the provided incentive is very low).

The cost of the effort increases as the sum of the effort exerted in activity X and Y increases and it does not depend on the worker's ability in performing activity X (i.e. $c(e) = c(e_x + e_y)$ for both worker H and L). The effort cost function is convex and its outcomes and the associated payoffs for the worker are displayed in Table 2.1 (the cost function is the same used in Brown *et al.* (2004) with the only difference that in this paper the agent only exerts effort in one activity). The worker who has not been chosen receives an unemployment benefit of 10 ECUs.

---

[2] Since each session is played by 28/32 participants, the probability to meet the same group of players during the second part of the game is quite low; moreover, participants have no possibility to communicate with each other and thus to recognize players they have already been matched with.

[3] At the end of part one the subjects did not receive any feedback information about their payoff or the strategies of the other players and this was common knowledge.

[4] We employed a conversion rate of 10 ECUs = €1.

Table 2.1. Payoff for the selected worker.

$e_x + e_y$	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Earnings for the worker =(W- Ce)	50	49	48	46	44	42	40	38	35	32

The principal's payoff entirely depends on the chosen level of effort by the hired worker in activity X. In every period of the game he is endowed with a budget E equals to 100 ECUs, which is entirely spent in order to pay both the wage of the agent and of the hired worker. The principal gets the 85% of the value produced by the hired worker in activity X. Depending on which worker is hired, the principal's payoff differs according to the profit's scheme illustrated in Table 2.2, which is calculated as follows:

$$P_p^k = (e_x + \alpha_k) \cdot 8.5 \cdot RV$$

Where, the productivity factor α indicates the worker's ability with $k = \{H, L\}$ such that $\alpha_H > \alpha_L$ and $P_p^H > P_p^L$; in particular, in our experiment $\alpha_H = 0.5$ and $\alpha_L = 0$.

The Random Value RV can affects both positively or negatively the surplus generated by activity X, $RV \in \left[\underline{\upsilon}, \overline{\upsilon}\right]$ with $\underline{\upsilon} < 1$, $\overline{\upsilon} > 1$ and $E(RV) = 1$. The random component makes thus difficult for the Principal to infer which effort level has been chosen by the hired worker.

Table 2.2. Principal's payoff (when considering the RV affects the payoff with its expected value equals to 1).

Effort in X	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Value if H is hired	12.8	21.3	29.8	38.3	46.8	55.3	63.8	72.3	80.8	89.3
Value if L is hired	8.5	17.0	25.5	34.0	42.5	51.0	59.5	68.0	76.5	85.0

The agent receives a fixed wage w_m equals to 50 ECUs. As for the principal, her payoff increases depending on the exerted level of effort in activity X and in activity Y, according to the following rule:

$$P_M^k = w_m + ((e_x + \alpha_k) \cdot 1.5) + (e_y \cdot 4.3)$$

In the *Baseline and Selection* (*Profit Sharing*) *treatments*, the agent gets the 15% (7.5%) of the value produced in activity X and, as for the principal, depending on which worker is hired, the payoff coming from activity X differs, with the value produced by worker H being higher than the value produced by worker L, when considering the same level of effort (i.e. $\alpha_H > \alpha_L$ with $\alpha_H = 0.5$ and $\alpha_L = 0$). Differently, the profit produced in activity Y is not affected by the worker's ability. Most importantly, as easily notable in Table 2.3, when considering the same level of effort, activity Y is always less efficient than activity X in producing earnings, independently on whether the worker is of type H or L.

Table 2.3a. Agent's payoff depending on the effort exerted by the hired worker in activity X.

Effort X	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Value if L is hired	2.3	3.8	5.3	6.8	8.3	9.8	11.3	12.8	14.3	15.8
Value if H is hired	1.5	3.0	4.5	6.0	7.5	9.0	10.5	12.0	13.5	15.0

Table 2.3b. Agent's payoff depending on the effort exerted by the hired worker in activity Y.

Effort in Y	**0**	**1**	**2**	**3**	**4**	**5**
Produced value	0.0	4.3	8.5	12.8	17.0	21.3

2.3.2 Feedback

Information about part 1 is only provided at the end of the experiment. At the end of each period of part 2, we provide the principal with an imperfect feedback regarding the chosen level of effort by the hired worker in activity X: indeed, he only receives information regarding the earnings produced in this activity, that depend both on the exerted effort and on a an unknown random component with expected value equals to one. Differently, the agent receives information about the exerted level of effort by the selected worker in both activity X and Y. Finally, both the agent and the principal are informed about the average effort chosen in activities X and Y by employees of type H and L which have been hired by other firms in that session and about how many employees of type H and L have been hired within the precedent period of the session[5]. We believe that spreading the information regarding other's behaviour in the game is critical in explaining individual's behaviour. Indeed, as previous research has stated (Kees *et al.*, 2008; Diekman *et al.*, 2011; Gino *et al.*, 2011), other's social norm's violation affects individuals choices: we thus expect that when a "corruptive" social norm begins to spread among players (i.e. the proportion of agents choosing the ex ante low ability worker increases, as well as the proportion of them exerting higher effort than ex ante high ability workers in activity Y), then the propensity of agents pursuing their personal interest at the expenses of those of the organization increases, generating a snowball effect (Chang and Lai, 2002).

2.3.3 Procedures

The experiment was programmed using zTree (Fischbacher, 2007). We conducted 21 experimental sessions at the experimental laboratory of the Max Planck Institute of Economics (Jena, Germany), from November 2013 to February 2013. Respectively, 216, 212 and 216 subjects participated to the Baseline, to the Selection and the Profit Sharing treatment, with about 28/32 individuals taking part in each session. The subjects were undergraduate students from the Friedrich Schiller University Jena recruited via the ORSEE software (Greiner, 2004).

Table 2.4. Participants and Treatments

Treatment	Participants	Group	Sessions
BSL	216	54	7
SEL	212	53	7
PS	216	54	7
Total	644	161	21

[5] Participants receive a similar feedback regarding others' behavior in part one only after part two is completed, and this is common information. This design assures that subjects play a one shot game in the first part of the game.

Once arrived in the laboratory, each participant was randomly assigned to one visually isolated computer terminal. It was common knowledge that the experiment was composed by two parts. Each subject received written instructions for the first part of the game. The instructions were read aloud and then individuals were asked to answer a set of control questions on the screen. Roles were then randomly assigned to subjects, who played the first part of the game (i.e. the one shot decision). Once having completed part one, instructions about part two of the experiment were distributed and read aloud. In the second part of the experiment, participants played the game 15 periods repeatedly, with a random re-matching of groups after each period.
At the end of each session, one period of part two was randomly extracted for payment. The payoff of part one then was added up and the sum was converted into Euro. The duration of each session was about 110 minutes and the average payment was 16 Euro, including a show up fee of 4 Euros.

2.4 Results

In the following sections we are referring to the *Baseline, Selection* and *Profit Sharing treatments* respectively as BSL, SEL and PS. First, in section 2.4.1, we analyze the agent's hiring decision comparing the SEL and PS treatments, where a conflict of interest between the principal and the agent is present, to the BSL treatment where, by design, there is no conflict of interest. Then, in section 2.4.2 we analyze the effort's suggestion made by agents to the selected workers. Moreover, we investigate the effort exertion by the hired workers in the three treatments, focusing on the relationship between effort in activity X (which is productive for the organization) and in activity Y (i.e. effort distorted in favor of the agent). Finally, in section 2.4.3, we provide the results regarding the profits both of the agents and of the principals, comparing the BSL, SEL and PS treatments.
Throughout the analysis, we will first present the results of the one shot decision taken in part 1, then we will focus on the dynamics observed over the 15 periods in part 2, when information about others' behavior is spread among participants[6].

[6] We analyze the results of the second part of the game considering each session as a single observation in order to take into account that spreading information regarding other participants' behavior may render participants' choices not to be independent within the same session.

2.4.1 Selection

In order to determine if the conflict of interests between the principal and the agent impacts the workers' selection process, we proceed in two steps. First we consider the selection in our BSL treatment where, by design, there is no conflict of interest. Then, we analyze the agents' choices in treatments SEL and PS, given the principals' suggestion. It may be the case, in fact, that a principal prefers to hire the L worker believing that, once hired he would exert higher effort in activity X than the H worker, because feeling less entitled to get the job, as shown by Montinari *et al.* (2012). Remember that in our design, only the agent is informed about the principal suggestion, but not the workers.

The hiring decision made by the principal in the BSL treatment as well as the suggestion sent to the agents in the SEL and PS treatments are reported in Table 2.5. In the BSL treatment, a fraction of principals hire the L workers and, similarly, in the SEL and PS treatments, they suggest the L worker to be hired by the agents.

Specifically, in part 1, we find that the proportion of the principals suggesting to hire a L candidate is not different in the BSL and SEL treatments, while it is slightly higher in the PS treatment than in the BSL one (two sample tests of proportion: BSL vs. PS: z=2.07, p=0.04; BSL vs. SEL: z=1.18, p=0.24; SEL vs. PS, z= 0.91, p=0.36).

Table 2.5. Fraction of L workers to be hired according to the suggestion of the Principal

	Part 1	Part 2			
		1-5	6-10	11-15	Overall
BSL	9.26%	29.03%	29.03%	26.84%	28.30%
SEL	16.98%	30.87%	38.32%	34.64%	34.61%
PS	24.07%	29.74%	34.54%	27.40%	30.56%

Note. In treatment BSL the suggestion of the principal is binding for the agents. In treatments SEL and PS the suggestion from the principal is cheap talk.

When looking at 15 repetitions in part 2, we find that the proportion of the principal suggesting to hire an L candidate is not different across treatments (Mann Whitney test: SEL vs. BSL: z= 1.02, p=0.31; SEL vs. PS: z= 0.51, p=0.61; BSL vs. PS, z= 0.57, p=0.56).

Consider now Table 2.6, which reports the hiring decision of the agents in the SEL and PS treatments depending on whether they follow or not the suggestion received by the principal.

Table 2.6. Percentage of agents NOT following the principals' suggestion

	Part 1	Part 2			
		1-5	**6-10**	**11-15**	**Overall**
SEL	28.30%	39.23%	41.33%	46.84%	42.46%
PS	24.07%	38.88%	41.48%	39.64%	40%

When the agents have the possibility to choose which worker to hire, a significant share of them decide to deviate from the suggestion of the principal. In the SEL and PS treatments, respectively 28.30% (N=15/53) and 24.07% (N=13/54) of agents in part 1 do not follow the suggestion of the principal. In part 2, overall the 15 periods, this happens in the 42.46% and in the 40% of the cases for treatment SEL and PS, respectively. Both in part 1 and 2, these proportions are not significantly different across the two treatments (part 1, two samples test of proportions: z=0.50, p=0.619; part 2, Mann Whitney test: z=0.83, p= 0.40). When restricting the attention only to those agents hiring the L workers, we find that in the SEL and PS treatments, in part 1, respectively the 87.50% (N=7/8) and 50.% (5/10) of agents select the L workers without following the principal's suggestion. In part 2, the 60.04% of agents hire the L worker ignoring the different suggestion of the principal, while in the PS treatment this happens the 60.60% of the times.

These first evidences show that agents deviate from the suggestion of the principal more often when hiring less able workers than when hiring more able ones. This can be considered as a first signal of a potential distortion on the hiring process operated by the presence of the conflict of interest.

Result 1. *In both the SEL and PS treatments, agents are more likely to hire low ability workers compared to the BSL treatment. In the SEL treatment, the percentage of low ability workers hired increases over time. The majority of low ability workers are hired without following the suggestion of the principal.*

Support for Result 1 can be found in Tables 2.5-2.7. Consider first the choices made in Part 1. In the Baseline treatment, 9.26% (N=5/54) of the principals decide to hire the less able workers, whereas a slightly higher proportion of agents enroll them in the SEL and PS treatments, 15.09% (N=8/53) and 18.52% (N=10/54), respectively. According to a set of two sample proportion tests, these differences across treatments are not statistically significant (BSL vs. SEL: $z = 0.92$ $p = 0.36$; BSL vs. PS: $z = 1.39$ $p = 0.164$; SEL vs. PS: $z = 0.47$ $p = 0.636$).

Table 2.7. Percentage of low ability workers hired

	Part 1	Part 2			
		1-5	**6-10**	**11-15**	**Overall**
BSL	9.26%	29.03%	29.03%	26.84%	28.30%
SEL	15.09%	39.08%	49.85%	54.03%	47.65%
PS	18.52%	43.72%	45.61 %	46.73%	45.36%

Note. In BSL treatment the principal decides which type of candidate to hire. In the SEL and PS treatments the agent decides which candidate hire after having received a non binding suggestion from the principal.

Consider now Part 2. Inspection of Table 2.7 reveals that, when considering the 15 periods overall, the percentage of low ability workers hired is significantly higher when there is a conflict of interests between the principal and the agent (Mann Whitney test: BSL vs. SEL: $z = 2.50$ $p=0.04$; BSL vs. PS: $z=2.23$ $p=.02$; SEL vs. PS: $z=0.12$ $p=0.90$). Moreover, when comparing the SEL and PS treatments, there are no significant differences in the hiring behavior of the agents. Note that, in the PS treatment, the agent receives a lower share of the profit compared to the SEL treatment, therefore, the incentives for acting in line with the organization's interest are lowered.

When analyzing the dynamics of the hiring decision across periods, it can be noted that in the SEL treatment the percentage of low ability workers hired increases of about 10 percentage points over time, from 39.08% in the first block of 5 periods to 54.03% in the last 5-periods block. While the percentage of low ability workers enrolled by the agent is progressively increasing in the SEL treatment, the hiring behavior of the principal remains almost constant in the PS and BSL treatments. Thus, the presence of conflict of interest has an effect on the hiring decision of agents. Since at the end of each period the information regarding the hired workers' average exerted effort both in activity X and Y is spread among participants, we expect the hiring decision, in particular in the SEL treatment, to be linked to workers' behavior. In the next section we thus analyze the exerted effort of workers in activity X and Y.

2.4.2 Effort

In the SEL and PS treatment, after the agent has chosen whether to hire the high or the low ability agent, she has the opportunity to suggest him which level of effort to exert in activity X and Y.

We observe that agents suggest workers which level of effort to exert in activity X and Y depending on their ability. In particular, in the second part of the SEL treatment, the agent is suggesting the low ability candidate to choose a level of effort in activity Y equals to 3.29, while asking for a significant lower level of effort (2.59) when hiring the high ability one (High vs. Low ability worker, Mann-Whitney test: $z = 2.24$ $p=0.02$). We observe the opposite result when considering activity X: the agent suggests the low ability worker to choose a lower level of effort than the high ability one (3.44 vs. 4.07), even if the difference is not significant (High vs. Low ability worker, Mann-Whitney test: $z = 1.21$ $p=0.22$). In the PS treatment the agents behave according to the same pattern than in the SEL treatment. However, while the suggested effort in activity Y to the high and low ability workers is not significantly different (High vs. Low ability worker, Mann-Whitney test: $z = 0.32$ $p=0.75$;), the agents are significantly more likely to suggest to high ability workers to work harder than low ability ones in activity X (High vs. Low ability worker, Mann-Whitney test: $z = 1.72$ $p=0.08$).

In the following subsections we focus on the effort choices made by the hired worker. Once selected, the worker simultaneously chooses the effort level to exert in activity X, which benefits both the principal and the agent, and in activity Y, which inefficiently only provide benefit to the agent.

When comparing the three treatments, we proceed as follows. First, we analyze the effort choices in activities X and Y, irrespectively of the workers' ability. Second, we concentrate on the differences in effort exertion both in activity X and Y, driven by differences in the ability of the selected workers (both within treatments and between treatments). Finally, we present the results of a set of multiple regression models estimating simultaneously the effort choices in activity X and Y.

2.4.2.1 Effort Exertion in Activity X

The effort exerted in activity X is profitable for the whole organization, i.e. both for the principal and the agent. As a consequence, a worker, when being concerned for the agent's welfare, does not necessarily have to exert effort in activity Y.

Both in the BSL and in the SEL treatments, there are no explicit incentives to make the workers to exert effort. In particular, since workers are paid a fixed wage and effort is costly to exert, a payoff maximizer worker chooses the minimum enforceable level of effort (i.e. 1) in activity X. Conversely, in the PS treatment, the workers receive, on the top of the flat wage, the 7.5% of the value produced in activity X. This incentive, however, is not high enough to change the optimal effort choice for a payoff maximizer worker, who should choose to exert the minimum enforceable effort level in the PS treatment too.

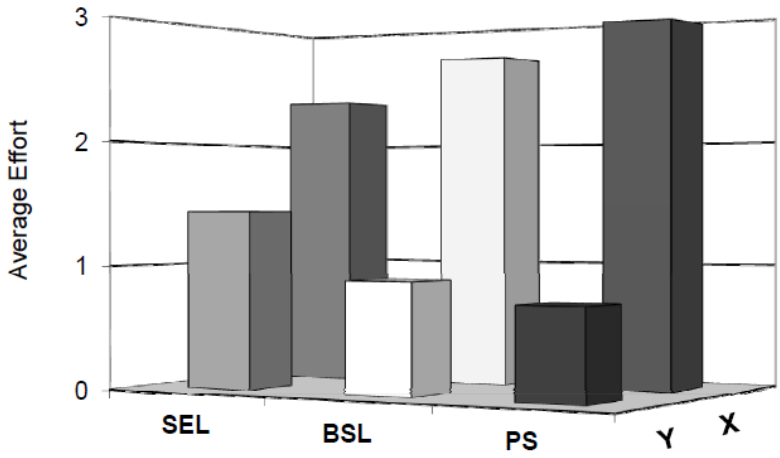
We find that, in both parts of the game and irrespectively of their abilities, workers exert an effort higher than the minimum enforceable one in activity X, as stated in Result 2.

Result 2. *In the absence of explicit incentives to make the hired worker exert a costly effort, in all treatments, workers choose to exert a positive effort in favor of the organization, higher than the minimum enforceable one. In the PS treatment, hired workers choose a higher effort level than in BSL and SEL treatments.*

Support for Result 2 can be found in Table 2.8 and Figure 2.1. In part 1, workers hired in the BSL and in the SEL treatments exert an average effort of 2.76 and 2.79, which are not significantly different from each other (Mann-Whitney test, $p=0.54$, $z= 0.61$). The average effort in the PS treatment, however, is equal to 3.66, which is significantly higher than both the BSL and SEL treatments (PS vs. BSL: Mann-Whitney test, $p=0.00$, $z= 2.60$; PS vs. SEL: Mann-Whitney test, $p=0.014$, $z= 2.47$).

In part 2, workers in the BSL and SEL treatments exert, over the 15 periods, an average effort of 2.62 and 2.30 respectively, which are not significantly different from each other (Mann-Whitney test, $p=0.11$, $z= 1.60$). When considering the average effort exerted in the PS treatment, it is significantly higher than in the SEL treatment but not statistically different than in the BSL treatment (PS vs. BSL: Mann-Whitney test, $p= 0.27$, $z= 1.09$; PS vs. SEL: Mann-Whitney test, $p=0.05$, $z= 1.91$).

Figure 2.1. Average effort exerted in activity X and Y in part 2, by treatment.



2.4.2.2 Effort Exertion in Activity Y

Consider now the effort exertion in activity Y, which only benefits the agent but not the organization. In our setting, exerting effort in activity Y represents an effort distortion in form of private benefit lavished to the agent who is responsible for the selection. When comparing the SEL and the BSL treatments, we can control for workers' distributional concerns towards the agent, since in the BSL treatment the latter has no power in the hiring decision. Result 3 summarizes our findings:

Result 3. *Workers significantly increase their effort level in activity Y in the SEL treatment compared to the BSL one. In the PS treatment, sharing a small part of the organization's profit with the workers significantly lowers the effort exerted in favor of the agent, to the level of the BSL treatment.*

Support for Result 3 can be found in Table 2.8 and Figure 2.1. In part 1, workers hired in the BSL treatment exert an average effort of 0.78 in activity Y, which is significantly lower than the average effort of 1.36 exerted in the SEL treatment (Mann-Whitney test, $p=0.01$, $z= 2.61$). In the PS treatment, however, the introduction of a small incentive for workers to work in activity X is sufficient in significantly reducing the effort exerted in activity Y. In particular, the chosen effort level in activity Y is not significantly different in the PS and BSL treatments. (Mann-Whitney test, PS vs. SEL: $p=0.00$, $z= 3.53$; PS vs. BSL: $z=0.84$, $p=0.40$).

In the second part, when considering the average effort exerted in activity Y over all the 15 periods, the results obtained in part 1 are confirmed: the average effort exerted in activity Y in the SEL treatment is equal to 1.43 and it is significantly higher than in the BSL treatment, where it is equal to 0.89, (Mann-Whitney test, BSL vs. SEL: z= 3.00, p=*0.00,*). When introducing a profit sharing however, the effort distortion in favor of the agent is significantly reduced to 0.73, a significantly lower level than the one observed in the SEL treatment (Mann-Whitney test, PS vs. BSL: z=1.34, p=0.18; PS vs. SEL: z=3.00, =0.00).

2.4.2.3 Productivity and Effort Exertion in Activities X and Y

In the following section we are now considering the effort exerted in activities X and Y, depending on the worker's ability. In the previous paragraphs we have shown that the presence of conflict of interests between the principal and the agent has an effect on the workers' selection process, with low ability workers being more likely to be hired in the SEL treatment. Moreover, we have found a significant increase of the level of effort exerted in activity Y in the SEL treatment with respect to the BSL one. Therefore, in this section we aim at understanding whether workers' effort choices both in activity X and Y depends on their ability. Result 4 summarizes our findings.

Result 4. *In the SEL treatment, while the H workers do not modify their effort exertion in activity X and Y with respect to the BSL treatment, the L workers increase their effort in activity Y . The introduction of a small profit sharing in the PS treatment eliminates the effort distortion by the L workers both by increasing their effort in X and reducing their effort in Y.*

Support for result 4 can be found in Table 2.8 and Figure 2.2. In the first part of the game the effort exerted in activity X both in the Baseline and in the Selection treatment does not differ depending on whether Worker L or H is hired (Baseline treatment, Mann-Whitney test, p=*0.507,* z= 0.664; Selection treatment, Mann-Whitney test p=*0.898*, z= 0.127).

Table 2.8. Average effort in Project X and Activity Y across repetitions (Standard Deviation in Parenthesis)

Part 1			Part 2							
			1-5		6-10		11-15		1-15	
Baseline treatment										
	X	Y	X	Y	X	Y	X	Y	X	Y
H	2.65	.86	2.70	.93	2.55	1.02	2.42	.73	2.54	.90
	(1.87)	(1.12)	(.53)	(.60)	(.48)	(.45)	(.39)	(.35)	(.34)	(.48)
L	3.80	0	2.58	.96	2.65	1.18	2.32	.65	2.66	.92
	(3.42)	(0)	(1.18)	(.46)	(.72)	(42)	(.89)	(53)	(.63)	(.35)
H+L	2.76	.78	2.64	0.95	2.60	1.10	2.37	.69	2.62	.91
	(2.04)	(.48)	(.88)	(.61)	(.59)	(.43)	(.69)	(.44)	(.51)	(.40)
Selection treatment										
	X	Y	X	Y	X	Y	X	Y	X	Y
H	2.82	1.31	2.52	1.06	2.26	1.01	2.12	.84	2.38	1.03
	(1.75)	(1.35)	(.84)	(.61)	(.27)	(.62)	(.46)	(.59)	(.58)	(.55)
L	2.63	1.63	2.33	1.54	2.17	1.98	2.21	1.52	2.21	1.73
	(1.41)	(1.41)	(.66)	(.15)	(.19)	(.44)	(.77)	(.48)	(.54)	(.29)
H+L	2.79	1.36	2.43	1.35	2.21	1.50	2.16	1.18	2.30	1.38
	(1.69)	(1.35)	(.73)	(.52)	(.59)	(.72)	(.61)	(.62)	(.55)	(.58)
Profit Sharing treatment										
	X	Y	X	Y	X	Y	X	Y	X	Y
H	3.40	.59	3.20	.60	2.80	.48	2.74	.61	2.93	.57
	(1.98)	(.87)	(.71)	(.18)	(.75)	(.17)	(.88)	(.42)	(.71)	(.40)
L	4.8	.4	3.36	.83	3.00	.87	2.53	.68	2.96	.76
	(1.40)	(.52)	(.79)	(.55)	(1.00)	(.63)	(.49)	(.42)	(.72)	(.26)
H+L	3.66	.55	3.26	.72	3	.68	2.72	.64	3.06	.68
	(1.95)	(.82)	(.75)	(.41)	(.88)	(.48)	(.71)	(.40)	(.71)	(.44)

When considering the PS treatment, however, we find that the L workers exert significantly higher effort in activity X than the H workers (BSL, Mann-Whitney test, $p=0.01$, $z= -2.42$).

Both the High and the Low ability workers are more likely to exert effort in activity Y in the SEL treatment than in the BSL and PS one.

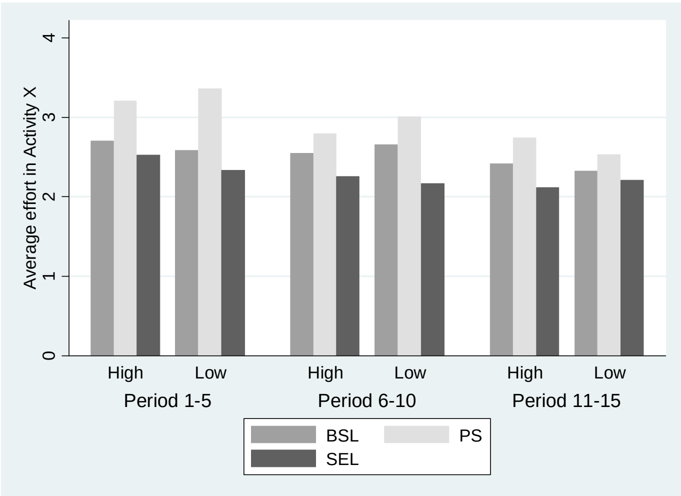
Now focus on Part 2 of the game. Table 2.8 shows that, in the first block of five periods of the SEL and BSL treatments, both H and L workers' effort in activity X does not differ neither when comparing workers within the same treatment (H vs. L workers, Mann-Whitney test: BSL, z=0.45, p=0.65; SEL, z=0.06 p=0.95), nor if considering each worker's effort across treatments (BSL vs. SEL, Mann-Whitney test: L workers, z=0.26, p=0.80; H workers, z=0.64 p=0.52). Moreover, such a behavioral pattern persists in the following blocks of periods. Differently, when considering the effort exerted in activity Y, only L workers distort their effort in the SEL treatment: in particular, while the H workers choose the same level of effort in both the BSL and SEL treatments, L workers are significantly more likely to work for the agent in the latter treatment in periods 1-5 (BSL vs. SEL, Mann Whitney test: L workers, z=2.62, p=0.00; H workers, z=0.57; p=0.56), in periods 6-10 (Mann Whitney test: L workers, z=2.68, p=0.00; H workers, z=0.06; p=0.94) and 11-15 (Mann Whitney test: L workers, z=2.49, p=0.01; H workers, z=0.38; p=0.70).

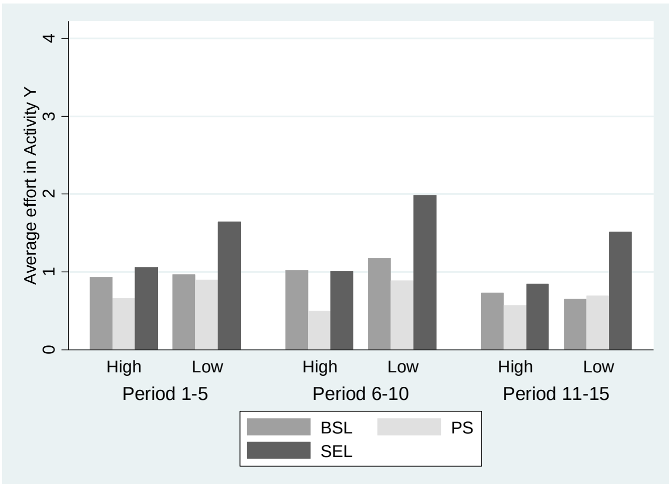
When considering the two types of worker within each treatment, we find that while the effort exerted in activity Y by low and high ability workers in the Baseline treatment is similar (Mann Whitney z=0.32 p=0.75, z=0.77; p=0.44, z=0.38 p=0.70, respectively for periods 1 to 5, 6 to 10 and 11 to 15), in the Selection treatment the low ability worker is more likely to work in favor of the agent than the high ability one (Mann Whitney test, z=2.05, p=0.04, z=2.49 p=0.01 and z=1.857 p=0.06 respectively for periods 1-5, 6-10 and 11-15).

In the PS treatment, when introducing a small profit sharing with the worker H workers do not change their effort exertion neither in activity X (Mann Whitney test PS vs. SEL z=1.47 p=0.14; BSL vs. PS z=0.96 p=0.34) nor in activity Y (Mann Whitney test PS vs. SEL and BSL vs. PS z=1.60 p=0.11); When considering L workers, the introduction of a profit sharing has a double positive effect both in the first and in the second part of the game: i) it increases significantly the effort exerted in activity X compared to the SEL treatment (Mann Whitney test, z=2.63, p=0.00 and z=2.10 p=0.03, respectively for the first and second part) bringing it back to a level which is not significantly different than in the BSL treatment (Mann Whitney test, z=0.75, p=0.45 and z=1.09 p=0.28, respectively for the first and second part); ii) it significantly reduces the effort devoted to activity Y in the SEL treatment to the level observed in the BSL one (PS vs. SEL, Mann Whitney test, z=2.08, p=0.04 and z=2.74 p=0.00, respectively for the first and second part)..

As an ulterior evidence of the distortive effect of conflict of interests in the organization, Figure 2.2 shows the different average levels of effort exerted in activity X and Y by H and L workers in the second part of the game, in the BSL, SEL and PS treatments.

Figure 2.2 Effort provided by the high and low ability workers in activity X (a) and Y (b) across treatments.





Finally, we present results from a set of regressions where we take into account the fact that workers' choice of X and Y are simultaneous. We proceed in two steps: in table 2.9 we present the results both of an OLS regression separate for X and Y and of a set of Zellner's seemingly unrelated regressions.

Models (1), (2), (4) and (5) report the results of the OLS regressions. Models (1) and (4) use as dependent variable the workers' choice of effort in activity X; models (2) and (5) instead considers as dependent variable the workers' choice of effort in activity Y. Models (3) and (6) reports the results of the Zellner's seemingly unrelated regression where respectively equations (1)-(2) and equations (4)-(5) are estimated simultaneously, accounting for the correlated errors at the same time.

As independent variables we include a set of dummies to identify the SEL and PS treatments, as well as a dummy to identify the ability of the worker ("L worker" which takes value 1 if the worker has low ability, and 0 otherwise). In models (3)-(6) we also include interactions between the treatment dummies and the worker's ability.

Table 2.9 reports the results of the Breusch-Pagan test of independence for models (3) and (6) which give us an estimate of the correlation between the errors of the two models. In both cases, the residuals from the two equations are not independent ($p<0.000$ in both cases), with a coefficient of correlation of the residuals around 10%.

By looking at model (3) it can be seen how the SEL and the PS treatments have two opposite effects on the effort exertion in X and Y compared to the BSL treatment. The SEL treatment has a positive and significant effect on the effort in activity Y while a negative and significant effect on effort in activity X. We interpret these effects as a signal of effort distortion due to the conflict of interest between the principal and the agent. The PS treatment has the opposite effect: positive and significant on the effort exerted in activity X and negative and significant for effort in activity Y.

The dummy accounting for the worker's ability is positive and significant for activity Y, indicating that once hired, the low ability workers exert higher effort in activity Y to reciprocate the choice of the agent.

Consider now model (6), where the interaction terms between the treatments and the worker type are included.

Table 2.9. The effort in Activity X and Y.

Model	**(1)**	**(2)**	**(3)**		**(4)**	**(5)**	**(6)**	
Estimation	**OLS**	**OLS**	**Seemingly unrelated regression**		**OLS**	**OLS**	**Seemingly unrelated regression**	
Dependent variable	**Effort x**	**Effort y**	**Effort x**	**Effort y**	**Effort x**	**Effort y**	**Effort x**	**Effort y**
Independent variables								
SEL	-.329*	.453***	-.308***	.462***	-.121	.271	-.097	.288***
	(.184)	(.119)	(.092)	(.058)	(.247)	(.219)	(.115)	(.073)
PS	.290	-.243*	.326***	-.242***	.387	-.274*	.418***	-.273***
	(.258)	(.133)	(.091)	(.058)	(.269)	(.148)	(.113)	(.072)
L worker	.075	.387***	.086	.370***	.390	.170	.415***	.149
	(.184)	(.123)	(.077)	(.049)	(.346)	(.167)	(.143)	(.090)
SEL x	-	-	-		-.566	.471	-.598***	.472***
L worker					(.468)	(.283)	(.193)	(.122)
PS x	-	-	-		-.330	.147	-.332*	.154
L worker					(.413)	(.250)	(.193)	(.122)
Period 6-10	-.265**	.081	-.296***	.090	-.256**	.073	-.287***	.081
	(.098)	(.052)	(.090)	(.057)	(.094)	(.053)	(.090)	(.057)
Period 11-15	-.436***	-.137**	-.467***	-.128**	-.422***	-.150**	-.452***	-.141**
	(.098)	(.063)	(.090)	(.057)	(.096)	(.062)	(.090)	(.057)
Constant	2.895***	.814***	2.90***	.808***	2.80***	.883***	2.80***	.876***
	(.137)	(.119)	(.082)	(.052)	(.153)	(.142)	(.089)	(.056)
N	2415	2415	2576	2576	2415	2415	2576	2576
Subjects								
Repetitions	1-15	1-15	1-15	1-15	1-15	1-15	1-15	1-15
Part 1	No	No	No	No	No	No	No	No
F/ Wald	(5,20)= 7.19***	(5,20)= 23.77***	-	-	(7,20)= 7.19***	(7,20)= 23.96***	-	-
R2	0.026	0.083	0.029***	0.083***	0.030***	0.089***	0.033***	0.088***
Breusch-Pagan test of independence	-	-	chi2(1) = 40.236***		-	-	chi2(1) = 43.667***	

In Models (1) (2) (4) and (5) Std. Err. adjusted for 21 clusters in session. In Models (3) and (6) standard errors are bootstrapped at the level of session.

Results are overall confirmed with small differences. The effect of the SEL treatment is now de-composed into two parts: there is a significant increase in the effort exerted in activity Y for both H and L agents. However, the negative and significant effect on the effort exerted in activity X is only driven by the L workers, as captured by the interaction term. Similarly, when looking at the effect of the PS treatment, we find a confirmation of the positive effect on X and of the negative effect on Y, as evidenced in model (3) but, in addition, it is associated with a positive effect on the effort exertion in activity Y by the L workers. Finally, L workers are more likely to exert effort in activity X, indicating that the L workers are in general more reciprocal than the H workers, this result can be explained by *induced* reciprocity, see

(Montinari et al., 2013), according to which the candidate who is less entitled for the position feels more in debt once hired and exert higher effort than the most entitled one.

2.4.3 The profit in the organization

An important question that we want to answer in this study is whether conflict of interests actually causes a welfare loss to the organization. In particular, in this section we want to analyze the impact of the above described distortions in the hiring process and of the effort provision by workers on the payoffs of each member of the organization. We first compare the profit made by the principals in the SEL and in the BSL treatments: while in the first part of the game there is no difference in the profit earned by the principals with respect to treatments, in the second part the principals' profits, independently on which worker is hired, decrease from an average of 25.89 ECUs in the BSL treatment to 22.22 ECUs in the SEL one (Mann-Whitney, z=1.21, p=0.22 in period 1 to 5, z=1.85 p=0.06 in period 6 to 10 and z=2.17 p=0.03 in periods 11 to 15).

On the other hand, when investigating the agents' profit, as resulted both from activity Y and activity X, we get the opposite result: their income is significantly higher in the SEL treatment than in the BSL one and this result is true both for the first (Mann-Whitney, z=2.09, p=0.04) and the second part (Mann-Whitney, z=2.24, p=0.02) of the experiment.

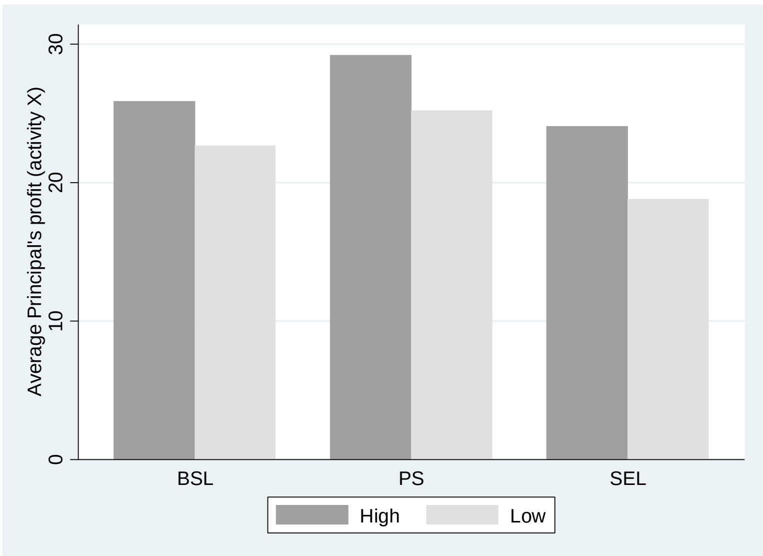
Result 5. *The presence of conflict of interest decreases the principal's profit while increasing the agent's one.*

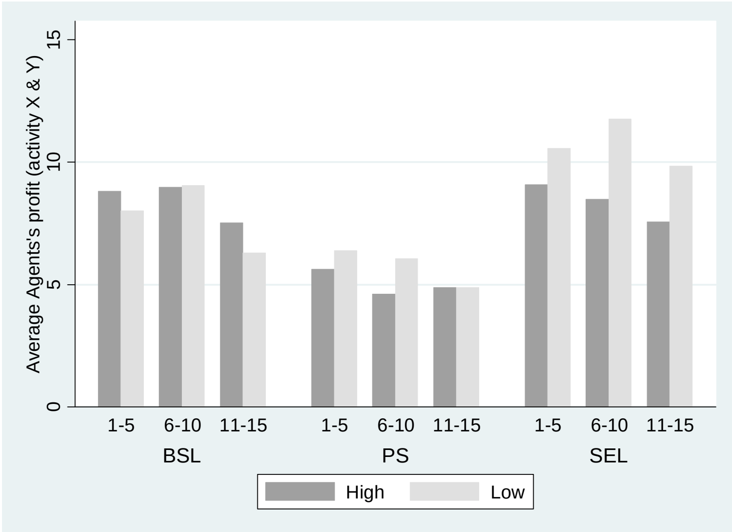
In order to further investigate whether the above described results depends on the distortion in the selection process, we study the profit provision to the principal and to the agent when the low (high) ability worker is hired, in the SEL treatment. Figure 2.3 shows that the agent's earnings are effectively higher when the low ability worker is hired, but this difference is never significant across periods. Even if the low ability worker is exerting a higher effort in activity Y than the high ability one (see Figure 2.2) and thus is producing higher earnings for the agent, both workers are devoting almost the same level of effort in activity X. In such a situation, the high ability worker is more productive and therefore balances the agent's profit.

We observe a similar result when considering the damage causes to the Principal by the distortion in the hiring process in the SEL treatment. While in the second part of the BSL treatment hiring the low ability worker instead of the high ability one decreases the principal's profit of 3.21 ECUs, when analyzing the profit loss in the SEL treatment then the effect is much stronger. In the SEL treatment, when agents hire the low ability workers instead of the high ability ones, the principals' profit decreases from an average of 24.05 ECUs to 18.80 ECUs. Figure 2.3 (a) shows that the profit loss remains almost constant in every block of periods even if not being significant (Mann- Whitney, z=1.60, p=0.11 and z=1.47 p=0.14 respectively for periods 1-5, 6-10 and 11-15).

When introducing a profit sharing with the workers, the principal's profits are restored to the same level of the BSL treatment (PS vs. BSL, second part, Mann- Whitney, z=0.64, p=0.52): in particular, even if the hiring distortion is still present, low ability workers exert their effort in activity X at a similar level than in the BSL treatment, instead of distorting their effort in favor of the agent. On the other hand, since in the PS treatment the agent only receives the 7.5% of the value produced in activity X by the hired worker, her profit are lower than in the BSL treatment (PS vs. BSL, second part, Mann- Whitney, z=2.62, p=0.01). Moreover, the hired workers are decreasing the effort exerted in activity Y, so that the agent's profits are significantly lower than in the SEL treatment (PS vs. SEL, second part, Mann- Whitney, z=3.13, p=0.00).

Figure 2.3. Earnings produced by the high and low productive workers in favor of the principal (a) and of the agent (b), in the second part of the BSL, SEL and PS treatments.





2.5 Conclusion

There is wide evidence that reciprocity is beneficial to organizations. Laboratory and field experiments have shown that workers are likely to provide more costly effort when the employer offers them higher wages than the minimum ones (see Henning-Schmidt et al. (2010) for an overview of recent results). However, some members of the organization may profit of such a mechanism for persecuting their personal illegitimate interests, on the detriment of the organization. In particular, we show that in a hierarchical organization, when the interests of delegated agents are not aligned with the principals' ones, then inefficiencies emerge in the hiring process of workers and on their productivity, once hired.

In our experiment the agent is delegated to select one worker out of two on the behalf of the Principal. Workers differ with respect to their relative ability so that one is more productive when exerting effort in favor of the organization. However, since the hired worker has the opportunity to exert effort just in favor of the agent, the latter has the incentive to disregard principal's expectations and hire the less able candidate. Low ability workers are indeed more likely to provide her with inefficient private benefits, because being aware of being less entitled to get the job and thus reciprocates the agent's choice, at the detriment of the organization. The presence of conflict of interests distorts the hiring process both in the *Selection* and in the *Profit Sharing* treatments, since a higher number of low abilities agents are enrolled in the firm. However, such a distortion negatively affects the organization only when workers are not involved in the organization through the profit sharing compensation scheme.

Our results suggest that sharing (a little part of the) profit both with workers and agents, rather than only with the latter ones, not only increases the motivation of low ability workers to exert effort in favor of the organization, but also prevent agents to take illegitimate advantage of their powerful position. Rewards system does not only represent a monetary motivational factor, but gives workers a cue to interpret the corporate culture of the organization. In particular, in our experiment, the profit's proportion shared with the workers is not high enough to push them to work harder, but enhance their feeling of membership to the organization as a whole.

In our experiment, there is no room for reputation building since workers, principals and agents play the game repeatedly with different group members. Hence, a relevant extension of our experiment would be to analyze whether the same results hold when allowing for long-run relationships between workers, agents and principals.

References

Akerlof, G. Labor Contracts as Partial Gift-Exchange. *Quarterly Journal of Economics* **97**, 543-569 (1982).

Akerlof, G., and Yellen, J. The Fair Wage-Effort Hypothesis and Unemployment. *Quarterly Journal of Economics* **105**, 255-283 (1992).

Bandiera, O., Barankai, I., and Rasul, I.. Social Connections and Incentives in the Workplace: Evidence from Personnel Data. *Econometrica* **77**, 1047-1094 (2009).

Bellemare, C., and Shearer, B. Gift Giving and Worker Productivity: Evidence From a Firm-Level Experiment. *Games and Economic Behavior* **67**, 233-244 (2009).

Brown, M., Falk, A., and Fehr, E. Relational Contracts and the Nature of Market Interactions. *Econometrica* **72**, 747-780 (2004).

Cain, D.M., Loewenstein, G., and Moore, D.A. Perverse effects of disclosing conflicts of interest. The Journal of Legal Studies **34**, 1-25 (2005).

Chang, J.J., and Lai, C.C.Is the Efficiency Wage Efficient? The Social Norm and Organizational Corruption. *Scandinavian Journal of Economics* **104**, 27-47 (2002).

Charness, G. Attribution and Reciprocity in an Experimental Labor Market. *Journal of Labor Economics* **22**, 665-688 (2004).

Charness, G., Cobo-Reyes, R., Jiménez, N., Lacomba, J.A., and Lagos, F. The Hidden Advantage of Delegation: Pareto Improvements in a Gift Exchange Game. *American Economic Review* **102**, 2358-79 (2012).

Dana, J., and Loewenstein, G. A Social Science Perspective on Gifts to Physicians from Industry. *Journal of the American Medical Association* **290**, 252–55 (2003).

Diekmann, A., Przepiorka, W., and Rauhut, H. Lifting the veil of ignorance: An experiment on the contagiousness of norm violations. *Discussion Papers* University of Oxford, Nuffield College, 2011.

Fehr, E., and Falk, A. Wage Rigidity in a Competitive Incomplete Contract Market. *Journal of Political Economy* **107**, 106-134 (1999).

Fehr, E., and Gachter, S. Reciprocity and economics: The economic implications of Homo Reciprocans. *European Economic Review* **42**, 845-859 (1998).

Fehr, E., Gachter, S., and Kirchsteiger, G. Reciprocity as a Contract Enforcement Device: Experimental Evidence. *Econometrica* **65**, 833-860 (1997).

Fehr, E., Kirchler, E, Weichbold, A., and Gächter, S. When Social Norms Overpower Competition: Gift Exchange in Experimental Labor Markets. *Journal of Labor Economics* **16**, 324-51 (1998).

Fehr, E., Kirchsteiger, G., and Riedl, A. Does Fairness Prevent Market Clearing? An Experimental Investigation.*The Quarterly Journal of Economics* **108**, 437-59 (1993).

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

Gino, F., Ayal, S., and Ariely, D. Contagion and Differentiation in Unethical Behavior. The Effect of One Bad Apple on the Barrel. *Psychological Science* **20**, 393-398 (2009).

Gintis, H. The Nature of the Labor Exchange of the Theory of Capitalist Production. *Review of Radical Political Economics* **8**, 35-54 (1976).

Gneezy, U., and List, J.A. Putting Behavioral Economics to Work: Testing for Gift Exchange in Labor Markets Using Field Experiments. *Econometrica* **74**,1365-1384 (2006).

Hannan, R. L., Kagel, J., and Moser, D.V. Partial Gift Exchange in an Experimental Labor Market: Impact of Subject Population Differences, Productivity Differences and Effort Requests on Behavior, *Journal of Labor Economics* **20**, 923-951 (2002).

Hennig-Schmidt, H., Rockenbach, B., and Sadrieh A. In search of workers' real effort reciprocity-a field and a laboratory experiment. *Journal of the European Economic Association* **8**, 817-837 (2010).

Jacquemet, N. Corruption as Betrayal: Experimental Evidence. *GATE Working paper 0506* (2012).

Kees, K., Lindenberg, S., and Steg, L. The Spreading of disorder. *Science* **322**, 1681-1685 (2008).

Kerr, J.L. and Slocum, J.W. Jr. Managing corporate culture through reward systems. *Academy of Management Executive* **1**, 99-107 (1987).

Koch, C., and + Schmidt., C. Disclosing conflicts of interest - do experience and reputation matter? *Accounting, Organizations and Society* **35**, 95-107 (2010).

Kramarz, F., and Thesmar, D. Social Networks in The Boardroom. *CEI Working Paper Series 2006-20*, Center for Economic Institutions, Institute of Economic Research, Hitotsubashi University, (2007).

Kube, S., Maréchal, M.A., and Puppe, C. Do Wage Cuts Damage Work Morale? Evidence From a Natural Field Experiment. *IEW Working Paper, (*2010).

Laffont, J., and Martimort, D. The Firm as a Multicontract Organization. *Journal of Economics & Management Strategy* **6**, 201-234 (1997).

Levine, D.K., Weinschelbaum, F., and Zurita, F. The brother in- law effect. *International Economic Review* **51**, 497–507 (2010).

Levitt, S.D., and Syverson, C. Market Distortions When Agents Are Better Informed: The Value of Information in Real Estate Transactions. *The Review of Economics and Statistics* **90**, 599-611 (2008).

Loewenstein, G., Cain, D., and Sah, S. The Limits of Transparency: Pitfalls and Potential of Disclosing Conflicts of Interest. *American Economic Review* **101**, 423-428 (2011).

Loewenstein, G., Sah, S. and Cain, D. The Unintended Consequences of Conflict of Interest Disclosure. *Journal of the American Medical Association* **307**, 669-670 (2012).

Maximiano, S, Sloof, R., and Sonnemans, J. Gift exchange and the separation of ownership and control. *Games and Economic Behavior* **77**, 41-60 (2013).

Mazar, N., Amir, O., and Ariely, D. The dishonesty of honest people: a theory of self-concept maintenance. *Journal of marketing research* **45**, 633-644 (2008).

Mehmet, B. Corruption and Supervision Costs in Hierarchies. *Journal of Comparative Economics* **22**, 99-118 (1996).

Montinari, N., Nicolò, A., and Oexl, R. Mediocrity and Induced Reciprocity. *Jena Economic Research Papers 2012-053*, Friedrich-Schiller University Jena, Max-Planck-Institute of Economics, (2012).

Norton, D., and Isaac R. *Experts with a conflict of interest: A source of ambiguity?*. Experimental economics. **15**, 260-277 (2012).

Prendergast C., and Topel, R.H. Favoritism in organizations. *Journal of Political Economy*, **104**, 958-78 (1996).

Robin, S., Rusinowska, A., and Villeval, M.C. Ingratiation and Favoritism: Experimental Evidence. *Documents de travail du Centre d'Economie de la Sorbonne 12032*, Université Panthéon-Sorbonne (Paris 1), Centre d'Economie de la Sorbonne (2012).

Thiele, V. Subjective Performance Evaluations, Collusion, and Organizational Design. *Journal of Law, Economics, & Organization* **29**, 35-59 (2013).

Tirole, J. Hierarchies and Bureaucraties: On the Role of Collusion in Organizations. *Journal of Law, Economics, and Organizations*, **2**, 181-214 (1986).

Appendix

A.1 Experimental instructions *(translated from German)*

Welcome! You are about to participate in an experiment funded by the Max Planck Institute of Economics. Please switch off your mobile and remain quiet. It is strictly forbidden to talk to the other participants. Whenever you have a question, please raise your hand and one of the experimenters will come to your aid.

You will receive 4.00 Euros for showing up on time and participate. Besides this, you can earn more. The show-up fee and any additional amounts of money you may earn will be paid to you in cash at the end of the experiment. Payments are carried out privately, i.e., the others will not see your earnings.

During the experiment we shall speak of ECUs (Experimental Currency Unit) rather than Euros. The conversion rate between them is 10 ECUs = 1 euro.

This means that for each ECU you earn you will receive 0.1 Euro.

To simplify, in the following we are only speaking of male participants. This is to be understanding gender neutral.

The experiment consists of two parts. The instructions for the first part follow on the next page. The instructions for the second part will be distributed after all participants have completed the first part.

Instructions are identical for all participants and we read them aloud such that you can verify this.

INSTRUCTION OF PART 1

1. GROUPS FORMATION

In this experiment you will be matched with three other participants to form a group of four persons. We will refer to **each group as an organization, and to the four group members as Manager, Selector, Employee A** and **Employee B**:

- with 1/4 probability you will be the Manager;
- with 1/4 probability you will be the Selector;
- with 1/4 probability you will be Employee A;
- with 1/4 probability you will be Employee B.

This means that each participant has the same probability to be hired as **Manager, Selector, Employee A** and **Employee B.** You will be assigned to a group and to a role entirely at random. The computer will inform you of your role before the decision-making part of the experiment begins. No one of the participants will know at any point of the experiment the identity of the other people in his group. Therefore, all decisions are made anonymously.

2. DECISIONS WITHIN A ORGANIZATION

The structure of the decision-making within each organization is as follows.

1. The **Manager:**

- manages the organization and employs the S**elector** and the **Employee,** paying their wages.
- suggests to the Selector one of the two Employees to be hired.

2. The **Selector:**

- is informed about the suggestion of the Manager.
- receives a wage from the Manager and
- [only in treatment Baseline (BSL): hires the employee (A or B) suggested by the Manager to work on a project which is beneficial for the organization.]
- [in all other treatments: hires one of the two employees (A or B) to work on a project which is beneficial for the organization. He is free to follow or not the suggestion].

3. The **hired employee:**

- learns that he has been hired,
- receives a fixed wage from the Manager to work on a project which is beneficial for the organization.

- chooses a level of costly **effort to exert.** He can exert effort in two activities: **project X** and **activity Y**.
    - **Project X produces earnings for the Organization (both the Manager and the Selector)**. The effort exerted by employee in project X has to be an integer number from 1 to 10 (included).
    - **Activity Y produces only earnings for the Selector but not for the organization**. The effort exerted by employee in activity Y has to be an integer number from 0 to 5 (included).
- **The sum of effort exerted in project X and activity Y determines the Total Effort** exerted by the employee. The total effort has to be an integer number **between 1 and 10**.

For a given level of effort, the earnings produced by the effort exerted in project X are always greater than the earnings produced by the effort exerted in activity Y.

4. The employee who is **not hired** receives an unemployment benefit of 10 ECUs**.**

3. DISTRIBUTION OF EARNINGS WITHIN A ORGANIZATION

Earnings within the Organization are determined according to the following rules.

3.1 Earnings for the Manager

The Manager has a budget of 100 ECUs. From this Endowment, the Manager pays

- the wage to the Selector (equal to 50 ECUs) and
- the wage to the Hired Employee (equal to 50 ECUs).

for a total of 100 ECUs.
Then, the Manager receives the 85% of the value produced in project X by the Hired Employee.

Earnings for the Manager	=	**Budget**	-	**Wage for the Selector** **Wage for the Employee**	+	85% **Value from project X**

Value from project X
The value produced by project X depends:
a) on the effort exerted by the hired Employee
b) on the influence of a random component.

a) Concerning the Employee,

- For each level of effort, employee A is more productive than Employee B;
- If Employee B exerts an effort (at least) one level higher than employee A, then B is more productive.

Table 1 reports the **value from project X for the manager** (in ECUs) for each exerted effort level in project X depending on the fact that Employee A or Employee B is hired. Please note that these values already represent the 85% of the total value produced in project X.

Table 1: **Project X**

Chosen effort	1	2	3	4	5	6	7	8	9	10
Value if Employee A is hired	12.8	21.3	29.8	38.3	46.8	55.3	63.8	72.3	80.8	89.3
Value if Employee B is hired	8.5	17.0	25.5	34.0	42.5	51.0	59.5	68.0	76.5	85.0

Note also that, irrespectively from the employee who is hired, the Manager's earnings increase with higher effort levels.

b) Concerning the Random component

- Once the hired Employee chooses the effort for Project X, the value generated by project X for the Manager is affected by a random component in the following way: a computerized random draw selects one of these 5 numbers: 0.8; 0.9; 1; 1.1; and 1.2.
- The value produced in project X is then multiplied by the drawn number and determines the earnings for the Manager.

All the 5 numbers are equally likely for both the Employees: this means that both Employee A and Employee B have the same probability to obtain, as random component, 0.8; 0.9; and so on…

Example 1. If Employees A is hired and he exerts an effort of 3 in the project X, the value generated is 29.8. Depending on the random draw, the Manager will obtain from Project X:

- 23.8 ECUs if 0.8 is drawn as random component;
- 29.8 ECUs if 1 is drawn as random component;
- 32.3 ECUs if 1.1 is drawn as random component and

Example 2. If Employees B is hired and he exerts an effort of 4 in the project X, the value generated is 34. Depending on the random draw, the Manager will obtain from project X:

- 30.6 ECUs if 0.9 is drawn as random component;
- 40.8 ECUs if 1.2 is drawn as a random component.

Information for the Manager

At the end of the experiment (i.e. after part 2 ends) the Manager will receive information about part 1. The Manager will be only informed about the final value of Project X and **he will not be informed about**

- [NOT in the baseline Whether employee A or B was hired by the Selector];
- The level of effort chosen by the hired employee and
- The random component randomly drawn.

3.2 Earnings for the Selector

The Selector receives

- A fixed wage by the Manager (equal to 50 ECUs) to hire one employee;
- the 15% [only in treatment Profit Sharing (PS): 7.5%] of the value produced in project X by the Hired Employee;
- the value produced in activity Y by the Hired Employee (if he does not refuse it).

Earnings for the Selector =	**Wage paid by the Manager**	**+Value from project X**	**+ Value from activity Y (if he does not refuse)**

Value from project X

The value produced by project X for the selector is calculated in the same way as for the manager and therefore it depends on the effort chosen by the hired employee and by the random component.

- For each level of effort, employee A is more productive than Employee B;
- if Employee B exerts an effort (at least) one level higher than employee A, then he is the more productive.

Table 2 reports the **value from project X for the selector** (in ECUs) for each chosen effort level in project X depending on the fact that Employee A or Employee B is hired. Please note that these numbers already represent the 15% [only in treatment Profit Sharing (PS): 7.5%] of the total value produced in project X.

Table 2: **Project X**

Chosen effort	1	2	3	4	5	6	7	8	9	10
Value if Employee A is hired	2.3	3.8	5.3	6.8	8.3	9.8	11.3	12.8	14.3	15.8
	1.1	1.9	2.6	3.4	4.1	4.9	5.6	6.4	7.1	7.9
Value if Employee B is hired	1.5	3.0	4.5	6.0	7.5	9.0	10.5	12.0	13.5	15.0
	0.8	1.5	2.3	3.0	3.8	4.5	5.3	6.0	6.8	7.5

Note also that, irrespectively from the employee who is hired, the Selector's earnings increase with higher effort levels.

Value from activity Y

While the effort in activity X benefit the organization (i.e. both the manager and the selector [only in treatment Profit Sharing (PS): and the hired employee]), activity Y constitutes a private benefit for the selector.

The value produced by activity Y does not depend on whether Employee A or Employee B is hired and it is reported in Table 3.

Table 3 **Activity Y**

Chosen effort level	0	1	2	3	4	5
Value	0.0	4.3	8.5	12.8	17.0	21.3

Please note that, irrespectively from which employee is hired, the selector's earnings increase with higher effort levels in the Activity Y.

For a given level of effort (i.e. 2),

- the value produced by the effort chosen in project X benefits the whole organization and
- it is always greater than the value produced by the effort chosen in activity Y and this is true no matter which number the computer draws (0,8; 0,9; etc.) to be multiplied by the value generated in project X.

Note that the selector can refuse the value (eventually) produced in activity Y. If this happens, then the value is devoted in favor of the Manager.

At the end of the experiment (i.e. after part 2) the Selector will receive information about part 1. The Selector will be informed about the value generated in activities X and Y by the hired employee in part 1.

3.3 Earnings for the hired employee

The Hired Employee receives

- a fixed wage by the Manager (equal to 50 ECUs) to work on a project which is beneficial for the organization;
- [only in treatment Profit Sharing (PS): the 7.5% of the value he produces in project X;]

His earnings are determined by the wage received minus the cost of the total effort he exerts [only in treatment Profit Sharing (PS): , plus the 7.5% of the value he produces in project X].

Earnings for the Hired Employee	=	**Wages for the Hired Employee**	-	**Cost of the Total Effort chosen**	+	**Value from Project X**

Once the Selector hires the employee, the hired employee chooses the level of costly effort to exert.
The hired Employee can exert effort in two activities: project X and activity Y. The sum of the two efforts in project X and activity Y determines the Total Effort, which has to be an **integer number from 1 to 10 (included)**:

- **Project X produces earnings for the Organization (both the Manager and the Selector [only in treatment Profit Sharing (PS): and the Employee])**. The effort chosen by employee in project X has to be an integer number from 1 to 10 (included).
- **Activity Y produces only earnings for the Selector but not for the organization**. The effort chosen by employee in activity Y has to be an integer number from 0 to 5 (included).

For a given level of effort, the earnings produced by the effort chosen in project X are always greater than the earnings produced by the effort chosen in activity Y.

For each level of total effort chosen, the earnings of the hired employee are shown in table 4. Please note that these numbers already represent the wage minus the cost of total effort.

Table 4

Total effort =(effort in project X + effort in activity Y)	1	2	3	4	5	6	7	8	9	10
Earnings for the hired employee =(wage-cost of total effort)	50	49	48	46	44	42	40	38	35	32

[only in treatment Profit Sharing (PS): To the earnings in table 4, the 7.5% of the total value produced in project X has to be added, as shown in Table 5.

Table 5: **Project X**

Chosen effort	1	2	3	4	5	6	7	8	9	10
Value if Employee A is hired	1.1	1.9	2.6	3.4	4.1	4.9	5.6	6.4	7.1	7.9
Value if Employee B is hired	0.8	1.5	2.3	3.0	3.8	4.5	5.3	6.0	6.8	7.5

For example,

- when the **hired employee** [only in treatment Profit Sharing (PS): is A and he] chooses to exert:
  - effort= 1 in working project X and
  - effort=0 in working activity Y

  he has a total effort of 1, and as it can be seen from the table, [only in treatment Profit Sharing (PS): tables 4 and 5, his earnings are 50 + 1.1= 51.1 ECUs] his earnings are 50 ECUs;

- when the **hired employee** [only in treatment Profit Sharing (PS): is B and he] chooses to exert:
  - effort=1 in working project X and
  - effort=1 in working activity Y

  he has a total effort of 2, and as it can be seen from the table, [only in treatment Profit Sharing (PS): tables 4 and 5, his earnings are 49 + 0.8= 49.8 ECUs] his earnings are 49 ECUs

Please note that the numbers used in all the examples were hired arbitrarily. They are not intended to suggest how you might decide.

1.4 Earnings for the non-hired employee

The employee who is **not hired** receives an unemployment benefit equals to **10 ECUs.**

[NOT in the Baseline (BSL):

4. COMMUNICATION

When making his choice, the **Selector** may send one message to the **employee he has hired**. In this message, he can ONLY indicate two numbers representing the suggested effort level he would like the Hired employee to choose both in project X and in activity Y. These numbers are not commitment for the employee who can, if hired, choose any effort level he wants in project X and activity Y. The hired employee reads the message before choosing the level of effort.]

WHAT HAPPENS NEXT?

Before the experiment starts, in next screens, you will be asked to answer a few questions.
When the experiment starts you will be informed about whether you are an Manager, a Selector or an Employee in this experiment. In case you are an employee, it will be specified whether you are employee A or employee B.

5. SUMMARY

We will now briefly summarize the content of the instructions you have just read.
At the beginning of the experiment you will be randomly matched with three other participants to form a group of four people and you will be randomly assigned a role within this group which we will call "organization". You will be the Manager, the Selector or Employee A or Employee B.
The structure of the decision-making within each organization is as follows.

- The **Manager**
  - employs the Selector and the **Employee** and pays them a wage.
  - He sends a suggestion to the Selector about who to hire;
- The **Selector**
  - reads the suggestion of the Manager
  - [not in Baseline (BSL): The Selector can send a message to the hired Employee: he can indicate two numbers representing the effort level he would like the Hired employee to choose both in project X and in activity Y.]
  - [not in Baseline (BSL): chooses to] hire **one** of the employees to work for the Organization.
- Next, the **hired employee**
  - learns that he has been hired,

- [not in Baseline (BSL): reads the message sent by the Selector and] receives the wage from the Manager
- and then he chooses a costly effort level. He can exert a maximum total effort of 10 in project X and in activity Y:
  - The effort chosen in project X can be an integer number between 1 and 10;
  - The effort chosen in activity Y can be an integer number between 0 and 5.
- The hired employee's earnings decrease with higher effort.

- The **non-hired employee** receives an unemployment benefit = 10 ECUs.

At the end of the experiment (after part 2 is finished), the Manager is informed about the value generated in project X and about the average effort exerted in project X and activity Y by hired employees of type A and B, respectively

The Selector is informed about the effort level chosen in project X and in activity Y by the hired employee and about the average effort chosen in project X and activity Y by hired employees of type A and B, respectively.

Please note that the decision project of this part of the experiment will be performed only once. Please, raise your hand if you have any questions.

Hypothetical examples for demonstration purposes

Please note that the numbers used in all the examples were hired arbitrarily. They are not intended to suggest how you might decide.

1. Assume that the Selector hires employee B.

The employee B chooses to exert effort= 4 in project X and effort=1 in activity Y.
Employee B's earnings:
The employee B receives a wage of 50 ECUs and chooses a total effort 4 (in project X) +1 (in activity Y)=5.
The total earnings of employee B are **44** [only in treatment Profit Sharing (PS): +3=**47]** ECUs.

This situation results in the following earnings for the other participants in the group:
Manager's earnings:
The Manager receives revenue from the effort of the employee B in project X (for sake of simplicity, suppose that the computer randomly draws number 1 to be multiplied by the value produced in Project X):
The total earnings of the Manager are 100-50-50+34**=34** ECUs.

Selector's earnings:
The Selector receives a wage of **50** ECUs.
The Selector receives revenues from the effort of the employee B in project X and in activity Y.
The total earnings of the Selector are 50+6+4,3=**60.3** ECUs [only in treatment Profit Sharing (PS): 50+3+4,3=**57.3** ECU].

Employee A's earnings: The employee A receives an unemployment benefit of 10 ECUs.

2. Assume that the Selector hires Employee A.

The employee A chooses the effort X = 3 in project X and effort Y = 1 in activity Y.
Employee A's earnings:
The employee A receives a wage of 50 ECUs and chooses a total effort 3 (in project X) +1 (in activity Y)=4. The total earnings of employee B are **46** [only in treatment Profit Sharing (PS): 2.6=**48.6**] ECUs.

This situation results in the following earnings for the other participants in the group
Manager's earnings:
The Manager receives revenue from the effort of the employee A in project X (for sake of simplicity, suppose that the computer randomly draws number 1 to be multiplied by the value produced in working Project X): The total earnings of the Manager are 100-50-50+29.8=**29.8** ECUs.

Selector's earnings:
The Selector receives a wage of **50** ECUs.
The Selector receives revenues from the effort of the employee A in project X and in activity Y: The total earnings of the Selector are 50+5.3+4.3=**59.6** ECUs [only in treatment Profit Sharing (PS): 50+2.6+4.3=**56.9** ECUs].

Employee B's earnings: The employee B receives an unemployment benefit of 10 ECUs.

INSTRUCTION OF PART 2

In this part you will face a situation similar to that encountered in the first part.

As before:

- you will be matched with three other participants to constitute a organization.
- **your role is the same as in part 1** (i.e., you will be the Manager, the Selector, the Employee A or the Employee B, if you previously were, respectively, the Manager, the Selector, the Employee A or the Employee B);

The structure of the decision-making within each organization is, as before:

- o The **Manager** employs the Selector and the Employee and pay them a fixed wage (50 ECUs); he sends a suggestion to the Selector about who to hire;
- o The **Selector** reads the suggestion sent by the Manager and [not in Baseline (BSL): chooses to] hire **one** of the employees to work for the Manager.
- o [not in Baseline (BSL): The **Selector** can send a message to the hired Employee**:** the Selector can indicate two numbers representing the effort level he would like the Hired employee to choose both in project X and in activity Y.]
- o Next, the **hired employee** learns that he has been **hired**, [not in Baseline (BSL): reads the message sent by the Selector], receives the wage from the Manager equal to 50 ECUs and then he chooses a costly effort level. He can exert a maximum total effort of 10 in project X and in activity Y:
  - ▪ The effort exerted in project X has to be an integer number between 1 and 10;
  - ▪ The effort exerted in activity Y has to be an integer number between 0 and 5.

  The hired employee's earnings decrease with higher effort.

- The **non-hired employee** receives an unemployment benefit = 10 ECUs.
- The earnings within the organization are divided as in part 1.

But now

- Part 2 consists of **30 rounds**.
- In every round the computer randomly rematches members of groups to form a new one (i.e., the three participants you will be matched with are different ones with respect to the previous round);
  - o You will maintain your role throughout all the 30 rounds;
  - o You will never be informed of the identity of the participants you will be matched with.
- After each round, the Selector will be informed:
  - o about his own payoff
  - o about the effort chosen in activities X and Y by the employee he has selected;

- About the average effort chosen in activities X and Y by employees of type A and B, respectively, and about how many employees of type A and B have been hired.
- After each round, the Manager will be informed:
  - About his own payoff
  - About the average effort exerted in activities X by employees of type A and B, respectively, and about how many employees of type A and B have been hired.
- After each round, the Employee will be informed:
  - about his own payoff;
  - about how many employees of type A and B have been hired and about the average effort exerted in activities X and Y by selected employees of his own type.

How we determine your earning in part 2

- Once part 2 is over, one participant will be randomly selected. This participant will determine which of the 30 rounds of part 2 will be paid, by making a random draw from the urn containing 30 balls (numbered 1 to 30).
- The earnings corresponding to this round will be converted to Euros and paid out in cash.
- The outcome of the urn-draw will apply to all the participants. Thus, only 1 of the X decisions you will make in this part will be paid out. **Obviously each decision has an equal chance of being used in the end. So, think carefully when making your choice in each round!**

Chapter 3

The wrong man for the job: biased beliefs and job mismatching

3.1 Introduction

There is strong evidence that women, while globally facing higher unemployment rates than men, seem also to be segregated in some segments of the labor market: they are underrepresented in managerial and legislative occupations and over-represented in midskill occupations[1] (Bourmpoula et al., 2012). In this paper, we provide a theoretical foundation to explain the emergence of the gender gap and segregation in the job market, as a consequence of different levels of self-confidence of men and women, when abilities are equally distributed among them.

Different reasons have been found to explain the existence of the gender gap in the workplace. First, women may have innate lower (higher) abilities than men in some sectors and are thus less (more) likely to be selected when applying. However, even if the discussion about this topic is still open, recent research suggests that men and women do not differ much in their cognitive abilities and it is rather social and cultural factors that influence perceived or actual performance differences (Hyde, 2005; Spelke 2005). Second, women and men face a different trade-off when formulating their career and family plans, showing a link between relative wages and fertility (Erosa et al., 2002; Galor and Weil 1996). In particular, Dessy and

[1] The global female labor force was estimated to be 1.3 billion in 2012, — about 39.9 per cent of the total labor force of 3.3 billion.

Djebbary (2010) show that the shorter reproductive capability of women with respect to men cause them to be more constrained in their career-family choices, so that failure in coordination of women's marriage-timing decisions lead to persisting gender differences in career choices. Third, some studies have questioned the existence of a glass ceiling[2] (Cotter et al., 2001), which lead organization to discriminate women's promotion and thus prevent them to achieve the highest rank, when having equal abilities than men (Bassanini and Saint-Martin, 2008). However, a recent research by the Institute of Leadership & Management (2011), claims that women managers are rather impeded in their careers by lower ambitions and expectations, which lead them to a cautious approach to career opportunities, than by a glass ceiling. Recently, a similar possible explanation of gender segregation has been developed, which relies on different preferences of men and women regarding the job environments where they would like to work, ultimately affecting their job entry decisions. Laboratory (Gneezy *et al.,* 2003) and natural field experiments (Flory *et al.,* 2010) provide evidence of women being less likely to apply to competitive work-settings. In particular, this phenomenon seems to be associated with men being more confident than woman about their relative performance in a (mathematical) task (Niederle and Vesterlund, 2007). Whether women have lower self confidence than men is a long lasting question (Lenney, 1977)), which seems to be sustained by studies in social psychology (Furnham, 2001) showing females as less likely to perceive themselves as qualified to run for political office (Lawless and Fox, 2005), or expressing lower career-entry and career-peak pay expectations (Bylsma andMajor, 1992). In particular, when considering the job market, Barbulescu and Bidwell (2012) showed that, among MBA students, women's lower expectations of job offer success is one of the causes of their lower number of application to finance and consulting jobs with respect to men. Such expectations have not an empirical foundation since the authors found no evidence that women were less likely to receive job offers in any of these fields. Such a bias, however, lead women to accept lower salary offers than the ones accepted by their male counterparts (Bowles *et al.*, 2005). All this evidence suggests that self-confidence has a crucial role in explaining the observed gender gap in the workplace: in our study, differences in how women and men perceive themselves as having the abilities to fulfill the job responsibilities, with respect to other candidates, affect the

[2] Carol Hymowitz and Timothy D. Schellhardt were the first to use the term "glass ceiling" in their March 24, 1986 article in the Wall Street Journal, "The Glass Ceiling: Why Women Can't Seem to Break the Invisible Barrier That Blocks Them from the Top Job."

workers' application decision, with women being less likely to apply to the skilled segment of the market. We find that biased beliefs of women about their relative abilities with respect to men lead to an inefficient job matching equilibrium: low skilled but self-confident males are recruited in the skilled segment of the job market while high-skilled, but under-confident, women are relegated to the unskilled one.

We reasonably assume that abilities are equally distributed among men and women while underconfidence is not. As a consequence, the observed gender gap appears to be not (only) a problem of fairness, but a problem of efficiency. Since high-skilled women are less likely to apply for top positions, firms are thus selecting their workers not necessarily in a group containing the best fitted candidates. In our model, women self-select into job positions according to their (mis)perceptions about their opportunity to be successfully recruited. Moreover, once segregated in the unskilled segment of the labor market, women's wrong beliefs would not be disconfirmed by evidence, making the gender gap to persist over time. We thus sustain the thesis that segregation in the job application and hiring process is caused by (biased) workers' self-selection.

The unequal distribution of males and females in the (un)skilled sector is thus endogenous on workers' application choices, when sorting decision depends on different gender's misperception of relative abilities to fit the job. This job mismatching thus causes an efficiency loss in the job market, which can be canceled out by calibrating suitable affirmative actions that guarantee the participation of high skilled women to the job market.

During the past years, several policies have been proposed to establish gender (and minorities) equality in the job market (see Anderson (2004), for an historical view of affirmative actions). Affirmative actions, first instituted in US in the 1960s and 1970s by employers and educational institutions, are temporary measures designed to increase the employment and educational opportunities available to qualified women and other minorities by giving them preference in hiring, promotion and admission. Coate and Loury (1993) provide mixed results regarding whether affirmative actions eliminate negative stereotypes in the employer's beliefs. Affirmative actions, such as exogenous imposed quotas on the labor force composition, have been often criticized for being unfair and inefficient. Opponents to such a policy, claim that it is unfair to hire an individual for a job on anything other than his qualifications and skills. However, the problem relieved in this paper is that females do not even apply to such jobs, because mistakenly being aware of their relative lower ability with respect to other candidates.

Since applications to job is a time consuming process, they prefer to apply to lower-skilled jobs, thus avoiding the risk to be disregarded and to be unemployed when competing for a high skilled position. Indeed, in a recent laboratory experiment, Balafoutas and Sutter (2012) provide evidence that affirmative actions encourage women to enter competition more often, without negatively affecting efficiency. Niederle et al. (2013) obtained a similar result when experimentally testing the effect of quota in favor of women in competitive tournaments. Moreover, in contrast with a common critique to the implementation of quotas, they did not find a decrease in the minimum performance threshold when achieving a more diverse set of winners. In this study we thus provide a theoretical foundation of the positive effect of quotas in favor of women in the labor market. However, we are not claiming that affirmative actions, such as gender quota imposition, is the only solution to the problem. Our model suggests that the gender gap may be a result of a structural bias in the system, so that an effective desegregation policy should intervene early in life in order to provide educational programs designed to positively encourage the correct development of self-image in women and to promote new role models.

The role of self-confidence is of primarily importance in our life, having a realistic (and positive) view of ourselves and of our own abilities may increase our motivation, letting us to engage in what we really want and can do. On the other hand, a too much high level of confidence may push people to accept challenges or tasks that are not fitted for their true abilities, increasing their probability to fail when trying to reach their objectives. An under-confident person, conversely, may pass up opportunities or may decide not to try (hard enough) to reach his personal aims. Many studies in psychology and economics have analyzed the role of self confidence both from a practical and theoretical point of view. In particular, when considering the role played by self confidence in the labor market, most of the literature has been primarily focused in the agency model. In both Sautmann (2011) and Santos Pinto (2008; 2010) studies, the principal, who is aware of the agent's overconfidence[3], takes advantage of it by paying the worker a lower wage. Moreover, Bénabou and Tyrole (2002) analyze the role of self-confidence in influencing how people process information and make decisions in order to explain some "irrational" behaviors such as self-handicapping or self-deception. In the studies by Falk *et al.* (2006a, 2006b) and Andolfatto *et al.* (2009) self confidence is analyzed in a job

[3] Luis-Pinto (2008) refers to the term positive (negative) self-image as the agent's over(under) estimation of the productivity of effort.

searching framework. While Andolfatto *et al.* (2009) apply the ideas of Bénabou and Tyrole (2002) in a model of labor market search, Falk *et al.* (2006a) Falk et al. (2006b) show that wrong beliefs about relative ability affects unemployment duration, in turn determining worker's potential starting wages. When considering sorting decisions in the labor market and in educational attainments, in a recent study Filippin and Paccagnella (2012) provide evidence that the level of confidence of young agents may consistently affect their lives: they show that even small differences in initial confidence of people about their ability may lead to diverging patterns of human capital accumulation between otherwise identical individuals. Larkin and Leider (2012) and Dohmen and Falk (2011) experimentally demonstrate that different incentive schemes invite different employees to join the organization, depending on their behavioral biases. In a study closely relate to our topic, Santos Pinto (2012) analyze the emergence of the gender pay gap as a result of males and females different levels of self-confidence in the classic labor market signaling model by Spence (1973): overconfident men are more likely to invest in education than underconfident women, which in turn lead to a higher productivity of men with respect to women, thus generating a gender pay gap.

The reminder of this paper is organized as follows. Section 3.1 develops the theoretical model, Section 3.2 characterizes the equilibria both when workers are biased and not, while Section 3.3 discusses the results and concludes. Proofs of all results are in the Appendix.

3.2 The model

The job market environment is modeled in the following way: there are two heterogeneous firms *k* with $k \in \{G, B\}$ and two risk neutral workers 1 and 2. We identify worker 1 as being a woman and worker 2 as being a man. Workers are ranked according to their different relative abilities *i* with $i \in \{L, H\}$, and we assume that it is always possible to rank them. Workers have equal probability ($\delta = \frac{1}{2}$) to be ranked as having higher ($i = H$) or lower ($i = L$) ability than the other worker. We refer to *G* as the firm belonging to the skilled labor segment, which offers one job position, and to *B* as the firm belonging to the unskilled labor segment, which offers an unlimited number of job offers. Let θ_k^i be the ability of worker *i* in firm *k*, we assume that:

(2.1) $$\theta_B^H = \theta_B^L = \theta_B$$

(2.2) $$\theta_G^H > \theta_G^L > \theta_B$$

The time of the game is as follows: at $t = 0$, workers observe their type[4] and at time $t = 1$ they simultaneously apply to one job position. At time $t = 2$, if the employer has received at least one application, she hires one worker and pays him a fixed wage w_k^i, which is an increasing linear function of the expected ability in equilibrium $\hat{\theta}_k^i$, as showed in the following equation:

(2.3) $$w_k^i = \frac{\hat{\theta}_k^i}{2}$$

Unemployed workers and firms with vacant jobs get an outcome equals to zero. The employer maximizes her profit with respect to the hiring decision:

(2.4) $$\pi_k = \theta_k^i - w_k^i$$

Hypotheses (2.1) and (2.2) imply that both workers prefer to be hired by employer *G*. However, employer *G* prefers to appoint worker *H* to the less able one, independently on whether worker *H* is a woman or a man (respectively, worker 1 and 2)[5]. On the other hand, employer *B* is indifferent between the two workers. In particular, when receiving multiple applications, employer *B* hires both candidates while employer G observes with probability p the applicants' productivity and with probability $1-p$ she is not able to correctly infer candidates' characteristics. In the former situation, she hires the most preferred candidate while in the latter situation, she hires with equal probability one of the two candidates and pays him the average expected productivity. The unselected candidate will then be unemployed and get zero utility. Since there is no unemployment

benefit we assume that participation constraints are satisfied. We assume that wages are the result of a Nash bargaining between interested parts. Moreover, whether the employer has observed the candidates' ability or not is common knowledge (i.e. the employer cannot pretend not to have observed applicants' productivity when it was the case).

In the model, we assume that agents can only apply to one job position, not to both, at time $t = 1$. We consider it as a reasonable assumption because of the following reason: application to job positions is a time-consuming process. Applicants need to prepare a cover letter, ask for letters of recommendation, adjust their resume in order to present themselves as the ones

[4] When being unbiased, workers are able to correctly observe their relative ability at time $t=0$.

[5] In our model, we assume that employers do not have prior beliefs regarding gender ability, i.e. negative stereotypes. In particular, firms select workers only considering the applicants' abilities.

perfect fitting for the job, according to the firm's specific requirements. As a consequence, candidates do not apply to all the job offers in the market, but just select the ones that they think to be the more achievable from their own point of view. Finally, to define whether a position is achievable or not a candidate has thus to take into consideration many factors: the expected salary, the working environment, but also his (perceived) relative ability with respect to others competitors in the application process.

In equilibrium candidate *i*'s expected utility is then computed as follows, with $i \neq j$:

(2.5) $$EU_G^i = p(w_G^i) + (1-p)\left[\frac{1}{2}\left(\frac{\theta_G^i + \theta_G^j}{4}\right) + \frac{1}{2}(0)\right]$$

(2.6) $$EU_G^i = p(0) + (1-p)\left[\frac{1}{2}\left(\frac{\theta_G^i + \theta_G^j}{4}\right) + \frac{1}{2}(0)\right]$$

(2.7) $$EU_B = \frac{\theta_B}{2}$$

Candidate *i*'s utility thus depends on whether both agents apply to firm G and candidate *i* is the preferred one (2.5) or not (2.6). Equation (2.7) represents candidate *i*'s utility when applying to firm B.

3.3 The job matching equilibrium

In the next subsections we derive the job matching equilibrium both in the case where workers are able to perfectly observe their rank (subsection 3.3.1) and in the case where worker 1 has biased beliefs regarding his relative ability with respect to the other worker (subsection 3.3.2). In both subsections 3.3.1 and 3.3.2, we examine the job matching equilibria when considering different levels of transparency in the job market, so that the employer in the skilled job market segment is able to infer with probability *p* the workers' ability, and with respect to different wage differentials between the job positions.

3.3.1 The benchmark case

In this section, we examine the matching equilibria in the job market when candidates, at time *t= 0*, are perfectly able to observe whether they are ranked as the worker having the lowest or the highest ability in the market. This equilibrium will be used as a benchmark. In particular, we assume that it is always possible to rank candidates according to their abilities: each worker has equal probability $\delta = \frac{1}{2}$ of being an *H* type, with the other worker being an *L* type, or of being an *L* type, with the other worker being an *H* type. In particular agent *i*, when being unbiased and observing being of type *H (L),* assigns probability one to the other agent to be type *L(H).* First of all, we define the equilibria in the job market as being of four types, according to the efficiency of the matching between workers and firm, from a social planner point of view. In particular:

Definition 1 *We define efficient separating equilibrium the job matching equilibrium where worker H applies to firm G and worker L applies to firm B. On the contrary, we define inefficient separating equilibrium the job matching resulting in worker H applying to firm B and worker L applying to firm G. Similarly, the pooling equilibrium where both workers apply to firm G is defined as efficient whereas the opposite one as inefficient.*

The following proposition characterizes the set of equilibria in pure strategies.

Proposition 1 *Equilibrium in pure strategies. Suppose workers are unbiased relatively to their ranking position and the wages differential across job segments is high enough so that* $\theta_G^H + \theta_G^L - 4\theta_B \geq 0$*:*

- *If the job market is sufficiently transparent so that the employer is able to observe the applicants' productivity with probability* $1 \geq p > \overline{p}$*, with* $\overline{p} = \frac{\theta_G^H + \theta_G^L - 4\theta_B}{\theta_G^H + \theta_G^L}$*, then the job market equilibrium is the efficient separating one: type H applies to firm G and type L applies to firm B.*
- *If the market is characterized by a lower ability to recognize different applicants' abilities, so that* $p \leq \overline{p}$*, then there are not separating equilibria and the efficient pooling equilibrium emerges, where both agents apply to firm G.*

The intuition of the Proposition 1 is the following. Consider first the case when *p* is large. The strategy profile such that *H* type applies to *G* and *L* type applies to *B* is an equilibrium, because if worker *L* deviates, then *L* remains unemployed with high probability and therefore *L* has not incentive to deviate. There are not other equilibria because when *H* type is not matched with firm *G* then *H* has incentives to deviate since with high probability will be employed by firm *G* at a higher wage. On the other hand, if $p \leq \overline{p}$, then the efficient separating equilibrium does not survive anymore. In particular, since now the firms' ability to correctly discriminate between the candidates abilities is lower, worker L is now tempted to apply to firm G, having the opportunity to be hired and get $\frac{\theta_G^H + \theta_G^L}{4}$ with half probability.

It is straightforward to note that the efficient pooling equilibrium exists only if $\theta_G^H + \theta_G^L - 4\theta_B \geq 0$, otherwise lower skill workers are not willing to pay the "risk" of being unemployed to get a slightly higher wage and thus prefer to apply to firm B and get a salary for sure.

When $4\theta_B - \theta_G^H - \theta_G^L \geq 0$ the skilled and the unskilled labor segments are only slightly differentiated in terms of the respective salary levels. In such a situation we have that:

Proposition 2 *Equilibrium in pure strategies II. Suppose workers are unbiased relatively to their ranking position and the wages differential across job segments is such that* $\theta_G^H + \theta_G^L - 4\theta_B \leq 0$*:*

- *If* $p \leq \tilde{p}$ *with* $\tilde{p} = \frac{4\theta_B - \theta_G^H - \theta_G^L}{3\theta_G^H + \theta_G^L}$, *then both the efficient and inefficient separating equilibria exist.*
- *If* $p \geq \tilde{p}$ *then only the efficient separating equilibrium survives.*

Finally, it is easy to see that the inefficient pooling equilibrium where both workers apply to firm B does not exist in any conditions of the market (See the Appendix for the proof).

The above statements provide evidence that a job market characterized by firms which are capable of correctly evaluate candidates with respect to their relative abilities is the most efficient in matching workers and firms. Conversely, when firms are not able to differentiate candidates then the market ends up in a very inefficient equilibrium where low ability candidates get a job in the high-skilled segment whereas high ability candidates work in the low-skilled one. In such a framework, therefore, a social planner should just improve the transparency of the market in order to have workers to be selected on a meritocracy criterion.

3.3.2 Underconfidence and job matching

In this section we assume the following: worker 1, which we have identified as a woman, has now biased beliefs regarding her ranking position in the job market. In particular, worker 1, at time $t= 0$, when observing her assigned type, misperceives it and assigns higher probability of being the low ranked candidate than the actual one, depending on the strength of the bias ρ. ρ measures the woman's bias, that we define as underconfidence, with $0 \leq \rho \leq +\infty$.

When worker 1 is fully underconfident, $\rho \to +\infty$, she assigns probability equals to 1 to be an L type, independently on the actual ranking at time $t= 0$. Conversely, when $\rho = 0$, we are considering exactly the same situation of the benchmark. We assume that worker 2 is aware of the bias of worker 1 and can benefit of this bias. The following proposition looks at the extreme case when worker 1 is fully underconfident.

Proposition 3 *Suppose worker 1 is extremely underconfident, such that* $\rho \to +\infty$.

- *When the differential among salaries in the skilled and unskilled segment of the job market is* $\theta_G^H + \theta_G^L - 4\theta_B > 0$ *then, if* $1 \geq p > \overline{p}$, *with* $\overline{p} = \dfrac{\theta_G^H + \theta_G^L - 4\theta_B}{\theta_G^H + \theta_G^L}$, *worker 2 always applies to firm G while worker 1 applies to firm B, independently on their respective types.*
- *When* $\theta_G^H + \theta_G^L - 4\theta_B < 0$, *then the inefficient separating equilibrium always exists, independently on* ρ.

The intuition of Proposition (3) is as follows: since worker 1, because of the bias, assigns now probability 1 of being an L type, then she is cautious in taking the risk of remaining unemployed when applying to the skilled segment of the market. As a consequence, she prefers to apply to firm B. Worker 2, on the other hand, being aware of worker's 1 underconfidence, is now conscious that the competition in the skilled segment of the market is softer and thus applies to firm G, even when being an L type.

In the following Proposition we now generalize our results by analyzing which level of worker 1's bias makes the inefficient separating equilibrium to emerge.

Proposition 4 *Suppose that wages differential across job segments is such that* $\theta_G^H + \theta_G^L - 4\theta_B \geq 0$, *then if* $\rho \geq \widehat{\rho}$ *with* $\widehat{\rho} = \dfrac{4\theta_B - 4p\theta_G^H - (1-p)\left(\theta_G^H + \theta_G^L\right)}{(1-p)\left(\theta_G^H + \theta_G^L\right) - 4\theta_B}$ *we get the following equilibria:*

- *if* $p > \overline{p}$, *with* $\overline{p} = \dfrac{\theta_G^H + \theta_G^L - 4\theta_B}{\theta_G^H + \theta_G^L}$, *then the inefficient equilibrium emerges.*
- *If* $p \leq \overline{p}$, *then we have an efficient pooling equilibrium where both workers apply to firm G.*

On the other hand, when $\theta_G^H + \theta_G^L - 4\theta_B \leq 0$, *then if* $\rho \geq \breve{\rho}$ *with* $\breve{\rho} = \dfrac{4\theta_B - 4\theta_G^H}{(1-p)\left(\theta_G^H + \theta_G^L\right) - 4\theta}$, *worker 1 applies to firm B and worker 2 applies to firm G, independently on their types, and this is true for any value of p.*

Corollary 1 *When* $\rho \geq \breve{\rho}$ *then the probability of observing an efficient matching between firms and workers now decreases when the probability p to observe the applicants' abilities increases. When (female) workers are not biased, if* $p \geq \overline{p}$, *with* $\overline{p} = \dfrac{\theta_G^H + \theta_G^L - 4\theta_B}{\theta_G^H + \theta_G^L}$, *then the probability to observe an efficient matching is equals to 1 while when (female) workers are biased then it is equal to* $\dfrac{1}{2}$. *When* $p \leq \overline{p}$ *then it is equals to* $p\dfrac{1}{2}(1-p)$.

Since hiring the "wrong man for the job", just because biased but skilled women are not applying anymore, cause a higher welfare loss to the job market when wages are highly differentiated with respect to candidates' abilities than when $\theta_G^H + \theta_G^L - 4\theta_B \leq 0$, we focus our attention on the job market environment where hiring a low or a high ability candidate in the skilled segment of the market has a great impact on the performance, such that $\theta_G^H + \theta_G^L - 4\theta_B \geq 0$.

The intuition of Proposition 4 is thus the following. Consider worker 1 first: when being of type *H* but sufficiently biased, she now prefers to apply to the unskilled segment of the market, in particular when the probability that firms are able to distinguish candidates' abilities is high. This is because worker 1, who misperceives herself as being an *L* type, mistakenly assigns lower probability to be enrolled when competing with worker 2 in getting a job in firm *G*, which is able to observe relative abilities of candidates. The higher is *p*, then the higher is the probability that she is going to be unemployed when applying to the skilled segment of the market. She thus prefers to apply to firm *B* and get a salary equals to $\frac{\theta_B}{2}$ for sure. As a consequence, worker 1 will never become aware of her own bias, when segregated in the low skilled segment of the market. The market thus ends up with half probability in an efficient separating equilibrium since worker 2, who is self-confident, applies to firm *G* both when being an *H* and *L* type. Worker 2, who is aware of the bias of worker 1, faces now a lower probability to compete with worker 1 in the skilled segment of the job market and will thus be enrolled more often than in the benchmark. In particular, when the hiring decision of firm *G* is based on meritocracy ($p \geq \overline{p}$), the probability to get an inefficient job matching is maximized. In

such a situation, it would be better for firms to commit to choose by random, but such a commitment is not credible because once that more than one worker applies, firm has incentive to screen them to hire the best one.

3.3.3 Affirmative actions

In our model the ability of firms to screen workers is not sufficient to restore an efficient matching, because underconfidence prevents women to apply for the skilled segment of the market and therefore there is an inefficient self selection of candidates. The ability to screen workers is useless when there are no workers to screen. In this section we want to study which affirmative actions can be undertaken to restore an efficient matching.

Affirmative actions are usually designed to improve the employment or educational opportunities of individuals in disadvantaged group. These polices sought to eliminate the injustices so frequently associated with discrimination, but there is disagreement about how to design them, and the introduction of exogenous quota are particularly ostracized. Nevertheless, in our study we are presenting a theoretical explanation to introduce quota as a tool to close the gender gap in a job environment where discrimination is not at stake.

First, in our model, affirmative actions are not introduced to fight females' discrimination in the job market, since we assume that there are no negative stereotypes with respect to the gender of candidates. In particular, we are focusing on the supply-side of the labor market, claiming that affirmative actions would increase the pool of qualified applicants by offsetting those self-defeating beliefs that prevent women from achieving the career goals they have the abilities to reach.

Second, in this study we are focusing on gender quota, a very controversial type of affirmative action. Gender quotas have been criticized as a form of reverse discrimination (Glazer, 1975; Sher, 1975), which favor candidates with respect to other observable characteristics (i.e. gender) rather than merit. However, this is true only if we assume that abilities are not distributed equally among men and women. If you think that the gender gap you observe in reality actually reflects the gender gap in abilities, then introducing a quota to increase the number of females in the job market will effectively lead to a reverse discrimination and to an efficiency loss. On the other hand, if males and females have the same abilities but not the same self-confidence when evaluating them (Furnham, 2001; Kling *et al.,* 1999), then introducing an exogenously imposed quota would lead qualified but underconfident women to apply, increasing the diversity and efficiency of the labor force, without discriminating men. Recent studies have indeed supported the introduction of gender quota to increase women's willingness to compete (Villeval, 2012). Laboratory experiments (Balafoutas and Sutther, 2012; Niederle et al., 2013) have demonstrated that gender quotas do not result in less able women overtaking most able men. In the present study, besides providing a theoretical evidence which explains the emergence of the gender segregation in the job market as a result of women's underconfidence, we now demonstrate the efficacy of calibrated quota in closing the gender gap and restoring efficiency.

We thus define an affirmative action as an exogenous probability ϕ to hire a women independently on any evaluation of her skill. Clearly, if $\phi = 1$, the affirmative action turns out to be discriminatory because it prevents men to be hired. Such an extreme policy induces the same level of inefficiency observed in the job market when there is not any affirmative action, such that $\phi = 0$ and women are underconfident about their relative abilities.

The following proposition characterizes the optimal affirmative action ϕ that restores the efficient matching:

Proposition 5 *Suppose* $p \geq \overline{p}$*, and* $\rho \geq \hat{\rho}$*. If the affirmative action* $\phi \in \left[\frac{p\left[\rho\left(\theta_G^H+\theta_G^L\right)+\theta_G^L-3\theta_G^H\right]+\left(1+\rho\right)(4\theta_B-\theta_G^H-\theta_G^L)}{p\left[\rho\left(\theta_G^H+\theta_G^L\right)+\theta_G^L-3\theta_G^H\right]+\left(1+\rho\right)(\theta_G^H+\theta_G^L)}; \frac{4\theta_B-\left(1-p\right)(\theta_G^H+\theta_G^L)}{\left(1-p\right)(\theta_G^H+\theta_G^L)} \right]$ *is settled, then the efficient separating equilibrium is restored.*

Proposition (5) implies that a social planner, in order to maximize the probability to get an efficient matching between high-skilled firms and workers in the job market, should not just limit his intervention in improving the ability of firms to discriminate among agents' abilities. Indeed, when applying such a policy in an environment where a part of the workers (i.e. women) are underconfident when evaluating their relative ability, then the market ends up in an inefficient separating equilibrium, incurring in a high welfare loss. To restore efficiency, the social planner should thus impose an exogenous probability that assures the participation of high ability women to the skilled segment of the market, without impeding men's candidature, when being of type H.

3.4 Discussion

In this paper we have presented a stylized model which identifies in the bias of women regarding their relative abilities in the job market an explanation of the emergence of gender segregation. In a job market where abilities are distributed equally among women and men, we have shown that when the bias is relevant, then we get an inefficient matching between workers and firms: indeed, high abilities female workers, when being underconfident regarding their relative abilities, are not applying to the skilled segment of the market, and this result is exacerbated when the ability of the market to rank and select workers according to their abilities is higher. Self-confident males are then enrolled in the high skilled segment both when having high and low abilities, since the competition is softer. As a consequence, when women are biased, the probability to end up in an efficient matching between high skilled firms and high abilities workers is reduced. In this paper we provide a theoretical base to explain the importance of implementing affirmative actions to restore efficiency in the job market. Imposing a suitable quota to the participation of women in the high skilled segment incentivizes them to apply when having high ability (and being underconfident) but not when having low abilities. As a consequence, the quota does not prevent men to participate to the high skilled job market, when having high abilities, thus maximizing the efficiency of the job matching. We have shown that incentivizing women to participate in the high skilled segment of the

job market is not (only) a question of fairness.

Alternative explanations to the low proportion of women in many high-profile jobs include negative stereotypes, different family-career plans, or different abilities. More recently, an increasing number of studies (Gneezy *et al*., 2003; Niederle and Vesterlund, 2007; Datta Gupta *et al*., 2013) have advanced the hypothesis that women are less prone to enter into competitive environments. Moreover, a bunch of the literature has observed that women exhibit a lack of self-confidence in their own abilities compared to men (Kling et al., 1999; Lenney, 1977)). In this paper we thus provide a theoretical base to explain the emergence of gender segregation as a result of self selection of (underconfident) women in mediocre career choices where abilities are not a big issue and where competition is soft. In order to counterbalance the negative effect of the bias of women we thus sustain the importance of implementing calibrated affirmative actions, to induce them to participate to the skilled segment of the job market, even when mistakenly perceiving themselves as not having the competencies to well perform in such a job. Affirmative actions, even if receiving divergent attention, have been showed to incentivize women's participation without affecting efficiency of the market (Niederle *et al.*, 2013; Balafoutas and Sutther, 2012). However, more effective solutions may be implemented to recover the efficient matching between firms and workers in the long term. In particular, since the gender gap in self-confidence seems to develop early in life (Orenstein, 1994; Hoffman, 1972) and depending on factors such as socioeconomic environments and parental attitudes (Filippin and Paccagnella, 2012; Chowdry et al., 2011), our results suggest that a policy which intervenes to equally encourage the development of self-image in young women and men would be beneficial in improving the gender equality and the efficiency of the job market.

References

Anderson, T. H. *The pursuit of fairness: a history of affirmative action* (Oxford University Press, 2004).

Andolfatto, D., Mongrain, D., and Myers, G. Rational truth-avoidance and self-esteem. *Canadian Journal of Economics* **42**, 141-154 (2009).

Balafoutas, L. and Sutther, M. Affirmative action policies promote women and do not harm efficiency in the laboratory. *Science* **335**, 579-582 (2012).

Barbulescu, R. and Bidwell., M. Do women choose different jobs from men? mechanisms of application segregation in the market for managerial workers. *Organization Science* (2012).

Bassanini, A. and Saint-Martin, A. *The price of prejudice: Labour market discrimination on the grounds of gender and ethnicity* (Technical report, OECD, 2008).

Bourmpoula, E., Ekkehard, E., Mahmood, M., and Salazar-Xirinchs, J. M. *Global employment trends for women* (Technical report, International Labour Organization Geneva, 2012).

Bowles, H., Babcock, R., McGinn, L. C., and Kathleen, L. Constraints and triggers: Situational mechanics of gender in negotiation. *KSG Working Paper No. RWP05-051* (2005).

Bylsma, W. H. and Major, B. Two routes to eliminating gender differences in personal entitlement: Social comparisons and performance evaluations. *Psychology of Women Quarterly* **16**, 193-200 (1992).

Bénabou, R. and Tyrole, J. Self confidence and personal motivation. *The Quarterly Journal of Economics* **117**, 871-915 (2002)..

Chowdry, H., Crawford, C., and Goodman, A. The role of attitudes and behaviours in explaining socio-economic differences in attainment at age 16. *Longitudinal and Life Course Studies* **2**, 5-76 (2011).

Coate, S. and Loury, G. C. Will affirmative-action policies eliminate negative stereotypes? *The American Economic Review* **83**, 1220-1240 (1993).

Cotter, D. A., Hermsen, J. M., Ovadia, S., and Vanneman, R. The glass ceiling effect. *Social Forces* **80**, 655-681 (2001).

Datta Gupta, N., Poulsen, A., and Villeval, M. C. Gender matching and competitiveness: experimental evidence. *Economic Inquiry* **51**, 816-835 (2013).

Dessy, S. and Djebbary, H. High-powered careers and marriage: can women have it all. *The B.E. Journal of Economic Analysis and Policy: Advances in Economic Analysis and Policy* **10** (2010).

Dohmen, T. and Falk, A. Performance pay and multidimensional sorting: Productivity, preferences, and gender. *The American Economic Review* **101**, 556-590 (2011).

Erosa, A., Fuster, L., and Restuccia, D. Fertility decisions and gender differences in labor turnover, employment, and wages. *Review of Economic Dynamics* **5**, 856- 891 (2002).

Falk, A., Huffman, D., and Sunde, U. Do i have what it takes? Equilibrium search with type uncertainty and non-participation. *IZA DP working paper 2531* (2006a).

Falk, A., Huffman, D., and Sunde, U. Self-confidence and search. *IZA DP working paper 2525* (2006b).

Filippin, A. and Paccagnella, M. Family background, self-confidence and economic outcomes. *Economics of Education Review* **31**, 824-834 (2012).

Flory, J. A., Leibbrandt, A., and List, J. A. Do competitive work places deter female workers? a large-scale natural field experiment on gender differences in jobentry decisions. *NBER Working Paper No. 16546* (2010).

Furnham, A. Self-estimates of intelligence: culture and gender difference in self and other estimates of both general (g) and multiple intelligences. *Personality and Individual Differences* **31**, 1381-1405 (2001).

Galor, O. and Weil, D. N. The gender gap, fertility, and growth. *American Economic Review* **86**, 374-387 (1996).

Glazer, N. *Affirmative Discrimination: Ethnic Inequality and Public Policy (*Harvard University Press, 1975).

Gneezy, U., Niederle, M., and Rustichini, A. (2003). Performance in competitive environments: Gender differences. *The Quarterly Journal of Economics* **118**, 1049-1074 (2003).

Hoffman, L. W. Early childhood experiences and women's achievement motives. *Journal of Social Issues* **28**, 129-155 (1972)..

Hyde, J. The gender similarities hypothesis. *American Psychologist* **60**, 581-592 (2005).

Kling, K. C., Hyde, J. S., Showers, C. J., and Buswell, B. N. Gender differences in self-esteem: A meta-analysis. *Psychological Bulletin* **125**, 470-500 (1999).

Larkin, I. and Leider, S. L. Incentive schemes, sorting, and behavioral biases of employees: Experimental evidence. *American Economic Journal: Microeconomics* **4**, 184-214 (2012).

Lawless, J. L. and Fox, R. L. *It Takes a Candidate: Why Women Don't Run for Office* (Cambridge University Press, 2005).

Lenney, E. Women's self-confidence in achievement settings. *Psychological Bulletin* **84**, 1-13 (1977).

Niederle, M., Segal, C., and Vesterlund, L. How costly is diversity? Affirmative action in light of gender differences in competitiveness. *Management Science* **59**, 1-16 (2013).

Niederle, M. and Vesterlund, L. Do women shy away from competition? do men compete too much? *Quarterly Journal of Economics* **122,** 1067-1101 (2007).

Institute of Leadership & Management. *Ambition and gender at work* (Technical report, 2011).

Orenstein, P. SchoolGirls. Young Women, Self-Esteem, and the Confidence Gap. Bantam Doubleday Dell Publishing Group, Inc., 1540 Broadway, New York, NY 10036 (1994).

Santos Pinto, L. Positive self-image and incentives in organizations. *The Economic Journal* **118**, 1315-1332 (2008).

Santos Pinto, L. Positive self-image in tournaments. *International Economic Review* **51**, 475-496 (2010).

Santos Pinto, L. Labor market signaling and self-confidence: Wage compression and the gender pay gap. *Journal of Labor Economics* **30**, 873-914 (2012).

Sautmann, A. Contracts for agents with biased beliefs: Some theory and an experiment. *Working Papers 2011-10*, Brown University, Department of Economics (2011).

Sher, G. Justifying reverse discrimination in employment. *Philosophy and Public Affairs* **4(**2) (1975).

Spelke, E. S. Sex differences in intrinsic aptitude for mathematics and science?: A critical review. *American Psychologist* **60**, 950—958(2005).

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

Villeval, M. C. Ready, steady, compete. *Science* **335**, 544-545 (2012).

Appendix

A.1 Proof of proposition 1

By assumption $\theta_G^H + \theta_G^L - 4\theta_B \geq 0$. Let's first consider the case when $1 \geq p > \overline{p}$, with $\overline{p} = \frac{\theta_G^H + \theta_G^L - 4\theta_B}{\theta_G^H + \theta_G^L}$. We first show that there exists a separating equilibrium so that worker H applies to firm G and worker L applies to firm B. In particular we have that the following equations must hold:

(A.1) $$EU_G^H \geq EU_B^H \Rightarrow \frac{\theta_G^H}{2} \geq \frac{\theta_B}{2}$$

(A.2) $$EU_B^L \geq EU_G^L \Rightarrow \frac{\theta_B}{2} \geq p(0) + (1-p)\left[\frac{1}{2}\left(\frac{\theta_G^H + \theta_G^L}{4}\right) + \frac{1}{2}(0)\right]$$

Equations A.1 and A.2 determine the equilibrium payoffs respectively of worker H and L. First, it is trivial to observe that A.1 always holds because of assumption (2.2). Thus, the only profitable deviation from equilibrium is worker L applying to firm G. By rearranging (A.2) we obtain:

$$p \geq \frac{\theta_G^H + \theta_G^L - 4\theta_B}{\theta_G^H + \theta_G^L} = \overline{p}$$

where p is the probability that the firm is able to rank candidates according to their different abilities when receiving two applications. Therefore, worker L has no incentive to deviate from equilibrium when $p \geq \overline{p}$ with $\overline{p} = \frac{\theta_G^H + \theta_G^L - 4\theta_B}{\theta_G^H + \theta_G^L}$.

Conversely, according to Proposition 1, when $p \leq \overline{p}$ there is a pooling equilibrium where both workers apply to firm G. In particular, in equilibrium, the payoff of worker H and L are, respectively, the following:

(A.3) $$EU_G^H \geq EU_B^H \Rightarrow p\left(\frac{\theta_G^H}{2}\right) + (1-p)\left[\frac{1}{2}\left(\frac{\theta_G^H + \theta_G^L}{4}\right) + \frac{1}{2}(0)\right] \geq \frac{\theta_B}{2}$$

(A.4) $$EU_G^L \geq EU_B^L \Rightarrow p(0)+(1-p)\left[\frac{1}{2}\left(\frac{\theta_G^H+\theta_G^L}{4}\right)+\frac{1}{2}(0)\right] \geq \frac{\theta_B}{2}$$

Rearranging A.3, we obtain:

$$p \geq \frac{4\theta_B - \theta_G^H + \theta_G^L}{3\theta_G^H + \theta_G^L} = \tilde{p}$$

When A.3 holds, so that $p \geq \tilde{p}$, then worker H applies to firm G instead of applying to firm B and get $\frac{\theta_B}{2}$. From A.2 we conclude that the same result holds for worker L if $p \leq \frac{\theta_G^H + \theta_G^L - 4\theta_B}{\theta_G^H + \theta_G^L}$. It is straightforward to see that if $p \leq \overline{p}$ then A.3 always holds and thus any of the workers have an interest in deviating from the pooling equilibrium.

No other equilibria in pure strategies exist. Consider the inefficient separating equilibrium where worker L applies to G and worker H applies to B. Then it must hold that:

(A.5) $$EU_B^H \geq EU_G^H \Rightarrow \frac{\theta_B}{2} \geq p\left(\frac{\theta_G^H}{2}\right)+(1-p)\left[\frac{1}{2}\left(\frac{\theta_G^H+\theta_G^L}{4}\right)+\frac{1}{2}(0)\right]$$

(A.6) $$EU_G^L \geq EU_B^L \Rightarrow \frac{\theta_G^L}{2} \geq \frac{\theta_B}{2}$$

First, it is trivial to observe that equation A.6 always holds because of assumption (2.2). Thus, the only profitable deviation is worker H applying to firm G. Rearranging A.5 we have that this is true only if $p \leq \tilde{p}$, the opposite condition expressed in equation (A.3). Since we are assuming that $\theta_G^H + \theta_G^L - 4\theta_B \geq 0$, then $\tilde{p}$ is negative and thus it does not exist any positive p that makes this equilibrium to exist. Finally, consider the inefficient pooling equilibrium where both workers apply to firm B. The following equations (A.7) and (A.8), representing the payoff of worker H and L when both applying to B, should hold:

(A.7) $$EU_B^H \geq EU_G^H \Rightarrow \frac{\theta_B}{2} \geq p\left(\frac{\theta_G^H}{2}\right)+(1-p)\left(\frac{\theta_G^H+\theta_G^L}{4}\right)$$

(A.8) $$EU_B^L \geq EU_G^H \Rightarrow \frac{\theta_B}{2} \geq p\left(\frac{\theta_G^L}{2}\right)+(1-p)\left(\frac{\theta_G^H+\theta_G^L}{4}\right)$$

Worker H has always an incentive to deviate from equilibrium and apply to firm G since, rearranging (A.7), we get that worker H plays the inefficient pooling equilibrium only if *p* satisfies the following condition:

(A.9) $$p \le \frac{2\theta_B - \theta_G^H - \theta_G^L}{\theta_G^H - \theta_G^L}$$

However, there is no positive p that makes (A.7) hold. In particular assumption (2.2) implies that $\theta_G^H + \theta_B > \theta_B + \theta_B = 2\theta_B$. Since $2\theta_G^L > 2\theta_B$, we get that $\theta_G^H + \theta_G^L > 2\theta_B$ thus the numerator of (A.7) is negative while its denominator is positive.

A.2 Proof of Proposition 2

Assume that $\theta_G^H + \theta_G^L - 4\theta_B \le 0$. When the offered salaries across labor segments are not very different, the efficient separating equilibrium where worker H applies to firm G and worker L applies to firm B exists if both equations (A.1) and (A.2) hold. As we have seen in the previous paragraph, this is true when $p \ge \bar{p}$ with $\bar{p} = \frac{\theta_G^H + \theta_G^L - 4\theta_B}{\theta_G^H + \theta_G^L}$. Since we are assuming that $\theta_G^H + \theta_G^L - 4\theta_B \le 0$, such condition is always satisfied because $\bar{p}$ is a negative number and thus for any $p \ge 0$ the efficient separating equilibrium exists.

When $p \le \tilde{p}$, with $\tilde{p} = \frac{4\theta_B - \theta_G^H - \theta_G^L}{3\theta_G^H - \theta_G^L}$, it also exists the inefficient separating equilibrium such that worker H applies to firm B and worker L applies to firm G. According to this equilibrium equations (A.5) and (A.6) must hold, which is the case when $p \le \tilde{p}$, as we have seen when rearranging (A.5). Since we are assuming that $\theta_G^H + \theta_G^L - 4\theta_B \le 0$, the previous condition is satisfied and thus the efficient and inefficient separating equilibria coexist when $p \le \tilde{p}$.

The same arguments used in the proof of proposition 1 can be used to prove that the inefficient pooling equilibrium where both workers apply to B do not exist even when $\theta_G^H + \theta_G^L - 4\theta_B \le 0$. Moreover, as proved in Proposition 1 , the pooling equilibrium where both workers apply to G exists only if $p \le \bar{p}$, with $\bar{p} = \frac{\theta_G^H + \theta_G^L - 4\theta_B}{\theta_G^H + \theta_G^L}$. However, because we are assuming that

$\theta_G^H + \theta_G^L - 4\theta_B \le 0$ then $\overline{p}$ is negative and thus equation (A.4) never holds. As a consequence the efficient pooling equilibrium does not exist.

A.3 Proof of Proposition 3

Suppose $\rho \to +\infty$; worker 1 is fully underconfident and she always thinks to be an L type. As a consequence, since we assume that it is always possible to rank candidates, she considers worker 2 as being the higher ability candidate. First, we define $pr(\hat{H} \mid H) = \frac{1}{1+\rho}$ as the probability that the biased worker 1 assigns to be an H type when actually being so at time *t = 0*. Conversely, $pr(\hat{L} \mid H) = \frac{\rho}{1+\rho}$ represents the probability that worker 1 assigns to the fact of being a type L even when she is not. She then prefers to apply to the unskilled segment of the market, even when being an H type, when the following condition is satisfied:

(A.10) $$EU_{1,H}(G) \le EU_{1,H}(B) = \frac{1}{1+\rho}\left(\frac{\theta_G^H}{2}\right) + \frac{\rho}{1+\rho}\left(\frac{1-p}{2}\right)\left(\frac{\theta_G^H + \theta_G^L}{4}\right) \le \frac{\theta_B}{2}$$

When $\rho \to +\infty$, then worker 1 is strongly biased and always perceived herself as being the lower ability worker in the job market and thus $\frac{1}{1+\rho} \to 0$ and $\frac{\rho}{1+\rho} \to 1$. Rearranging inequality (A.10) we get that when $p \ge \overline{p}$ with $\overline{p} = \frac{\theta_G^H + \theta_G^L - 4\theta_B}{\theta_G^H + \theta_G^L}$, then the above condition is satisfied and worker 1 always prefer to apply to firm B, even when being an H type.

On the other hand, worker 2, independently on his type, prefers now to apply to firm G, since there is no risk of competing with worker 1 and thus get the following expected utility:

$$EU_2(G) \ge EU_1(B) = \rho\left(\frac{\theta_G^L}{2}\right) + (1-\rho)\left(\frac{1-p}{2}\right)\left(\frac{\theta_G^H + \theta_G^L}{4}\right) \ge \frac{\theta_B}{2}$$

It is trivial to observe that when the wage differential among the segments of the market is lower, such that $\theta_G^H + \theta_G^L - 4\theta_B \le 0$, then the inefficient separating equilibrium always exists, independently on *p*. On the other hand, when the opposite condition applies, such that

$\theta_G^H + \theta_G^L - 4\theta_B \geq 0$, when $p \leq \overline{p}$ then condition (A.10) is no more satisfied and worker 1, even if underconfident, is prone to "take the risk" of applying to firm G since the probability that the firm will be able to distinguish candidates's abilities is now low enough. In such a situation, then we result in a pooling efficient equilibrium where both workers apply to firm G.

A.4 Proof of Proposition 4

According to Proposition (4), when worker 1 is sufficiently underconfident, then an inefficient separating equilibrium emerges. First, we analyze the case when $\theta_G^H + \theta_G^L - 4\theta_B \geq 0$. From the analysis of Proposition (1) we have that, when female worker has no bias then we end up in a separating efficient equilibrium when $1 \geq p > \overline{p}$, with $\overline{p} = \frac{\theta_G^H + \theta_G^L - 4\theta_B}{\theta_G^H + \theta_G^L}$, and in a pooling efficient equilibrium when $p \leq \overline{p}$. We want to analyze when the worker 1's bias is relevant in determining the emergence of different equilibria.

We proceed by solving equation (A.11) with respect to ρ:

$$\text{(A.11)} \quad EU_{1,H}(G) = \frac{1}{1+\rho}\left[p\left(\frac{\theta_G^H}{2}\right) + \left(\frac{1-p}{2}\right)\left(\frac{\theta_G^H + \theta_G^L}{4}\right)\right] + \frac{\rho}{1+\rho}\left(\frac{1-p}{2}\right)\left(\frac{\theta_G^H + \theta_G^L}{4}\right) \leq \frac{\theta_B}{2}$$

With a little bit of algebra it yields:

$$\text{(A.12)} \qquad \rho \leq \frac{4\theta_B - 4p\theta_G^H - (1-p)\left(\theta_G^H + \theta_G^L\right)}{(1-p)\left(\theta_G^H + \theta_G^L\right) - 4\theta_B} = \hat{\rho}$$

Worker 1 thus prefers to deviate from the pooling equilibrium where both workers apply to firm G when ρ satisfies the above condition. To define the measure of the bias we have to consider the sign of the denominator in expression (A.12) which depends on p.

Depending on whether p is higher or lower than $\overline{p}$, we have that:

$$\text{(A.13a)} \qquad \text{if } \overline{p} \geq \frac{\theta_G^H + \theta_G^L - 4\theta_B}{\theta_G^H + \theta_G^L} \text{ the denominator of } \hat{\rho} \text{ is positive thus } \rho \geq \hat{\rho}$$

(A.13b) if $\overline{p} \leq \frac{\theta_G^H + \theta_G^L - 4\theta_B}{\theta_G^H + \theta_G^L}$ the denominator of $\hat{\rho}$ is positive thus $\rho \leq \hat{\rho}$

Since the numerator of (A.12) is always negative (because of (2.2) and because we are assuming that $\theta_G^H + \theta_G^L - 4\theta_B \geq 0$), when $p \leq \overline{p}$ we have that $\hat{\rho}$ is always negative so that condition (A.12) is never satisfied and worker 1 prefers to apply to firm G for any value of ρ. The intuition is as follows: when the probability that the job market is able to differentiate candidates with respect to their abilities is low enough, then worker 1, when being type H but underconfident, thinks to have opportunity to be enrolled when applying to firm G, even when competing with worker 2, since with probability p the firm chooses at random.
Conversely, when $p \geq \overline{p}$, then the probability that firm G observe candidates' relative abilities is too high and worker 1, when being underconfident in the measure of $\rho \geq \hat{\rho}$, prefers to apply to firm B. In such a situation worker 2 always applies to firm G, independently on his type, since he's aware that worker 1, because of the bias, will not apply and he will be enrolled for sure. Thus, conversely than in the benchmark, when $\rho \geq \hat{\rho}$ with $\hat{\rho} = \frac{4\theta_B - 4p\theta_G^H - (1-p)\left(\theta_G^H + \theta_G^L\right)}{(1-p)\left(\theta_G^H + \theta_G^L\right) - 4\theta_B}$ the job market results in an inefficient separating equilibrium when $p \geq \overline{p}$, rather than in an efficient separating one.
We now prove the existence of the inefficient separating equilibrium when the job market wages are so that $\theta_G^H + \theta_G^L - 4\theta_B \leq 0$. In Proposition (2) we have showed that when workers are not biased, then in equilibrium, worker H applies to firm G and worker L applies to firm B. Thus, to induce the inefficient separating equilibrium, the following condition must hold for worker 1, when being of type H and underconfident:

(A.14) $$EU_{1,H}(G) = \frac{1}{1+\rho}\left(\frac{\theta_G^H}{2}\right) + \frac{\rho}{1+\rho}\left(\frac{1-p}{2}\right)\left(\frac{\theta_G^H + \theta_G^L}{4}\right) \leq \frac{\theta_B}{2}$$

When solving it with respect to ρ we obtain:

(A.15) $$\rho \leq \frac{4\theta_B - 4\theta_G^H}{(1-p)\left(\theta_G^H + \theta_G^L\right) - 4\theta} = \breve{\rho}$$

As in the previous paragraph, the sign of the above inequality depends on the value of the denominator, which in turn depends on *p*. In particular, since we are assuming that $\theta_G^H + \theta_G^L - 4\theta_B \leq 0$ and because of (2.2), both the numerator and the denominator of (A.15) are negative. As a consequence, the inefficient separating equilibrium exists for any value of *p*, when worker 1 underconfidence is such that $\rho \geq \breve{\rho}$.

A.5 Proof of Proposition 5

In Preposition (4) we have showed that when $1 \geq p > \overline{p}$, with $\overline{p} = \frac{\theta_G^H + \theta_G^L - 4\theta_B}{\theta_G^H + \theta_G^L}$, then if worker 1 is sufficiently biased so that $\rho \geq \hat{\rho}$ with $\hat{\rho} = \frac{4\theta_B - 4p\theta_G^H - (1-p)\left(\theta_G^H + \theta_G^L\right)}{(1-p)\left(\theta_G^H + \theta_G^L\right) - 4\theta_B}$, then worker 1 applies to firm B and worker 2 applies to firm G, independently on their types. In order to restore the efficient matching of workers, we prove that the social planner should impose an exogenous probability ϕ which assure that, under these conditions, worker 1 will be employed in firm G regardless of her ability and will receive a wage equals to $\frac{\theta_G^H + \theta_G^L}{4}$:

(A.16)

$$EU_{1,H}(G) \geq EU_{1,H}(B) = \phi\left(\frac{\theta_G^H + \theta_G^L}{4}\right) + (1-\phi)\left\{\frac{1}{1+\rho}\left[p\left(\frac{\theta_G^H}{2}\right) + \left(\frac{1-p}{2}\right)\left(\frac{\theta_G^H + \theta_G^L}{4}\right)\right] + \frac{\rho}{1+\rho}\left(\frac{1-p}{2}\right)\left(\frac{\theta_G^H + \theta_G^L}{4}\right)\right\} \geq \frac{\theta_B}{2}$$

Rearranging inequality (A.16), we get:

(A.17)
$$\phi \geq \underline{\phi}\frac{p\left[\rho\left(\theta_G^H + \theta_G^L\right) + \theta_G^L - 3\theta_G^H\right] + (1+\rho)(4\theta_B - \theta_G^H - \theta_G^L)}{p\left[\rho\left(\theta_G^H + \theta_G^L\right) + \theta_G^L - 3\theta_G^H\right] + (1+\rho)(\theta_G^H + \theta_G^L)} = \underline{\phi}$$

When $\phi \geq \underline{\phi}$ worker 1 applies to firm G when being unbiased. In such a situation worker 2, when being an L type, does not take anymore advantage of worker 1's bias but prefers to apply to firm B. However, the exogenous probability ϕ has not to be too high, pushing worker 1 to apply to the skilled segment even when being an L type and thus preventing worker 2 to apply to firm G, when being an H type. In particular, we get an efficient separating equilibrium when worker 2 is an H type and worker 1 is an L type, only if worker 2 is not incentivize to apply to firm G also when being the low ability candidate so that we must have that:

(A.18) $$EU_{1,L}(G) \le EU_{1,L}(B) = \phi\left(\frac{\theta_G^H + \theta_G^L}{4}\right) + (1-\phi)\left[p(0) + \left(\frac{1-p}{2}\right)\left(\frac{\theta_G^H + \theta_G^L}{4}\right)\right] \le \frac{\theta_B}{2}$$

Rearranging (A.18), it yields:

(A.19) $$\phi \le \frac{4\theta_B - (1-p)(\theta_G^H + \theta_G^L)}{(1-p)(\theta_G^H + \theta_G^L)} = \overline{\phi}$$

When (A.17) holds, then high skilled worker 1 applies to firm G even if being biased while low ability worker 2 does not. Moreover (A.19) assures that worker 1, when being L type, applies to firm B so that worker 2 will not be prevented to enter the high skilled segment of the market, when being an H type. Under these conditions the probability of efficiency matching is maximized, even when worker 1 is biased.